


January 2012

A Location-Based Incentive Mechanism for Participatory Sensing Systems with Budget Constraints

Luis Gabriel Jaimes

University of South Florida, ljaimes@mail.usf.edu

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [American Studies Commons](#), and the [Artificial Intelligence and Robotics Commons](#)

Scholar Commons Citation

Jaimes, Luis Gabriel, "A Location-Based Incentive Mechanism for Participatory Sensing Systems with Budget Constraints" (2012).
Graduate Theses and Dissertations.
<http://scholarcommons.usf.edu/etd/4087>

This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.

A Location-Based Incentive Mechanism for Participatory Sensing Systems with Budget
Constraints

by

Luis G. Jaimes

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Computer Science
Department of Computer Science and Engineering
College of Engineering
University of South Florida

Major Professor: Miguel A. Labrador, Ph.D.
Wilfrido Moreno, Ph.D.
Sudeep Sarkar, Ph.D.

Date of Approval:
May 10, 2012

Keywords: Crowd-Sourced Sensing, Mobile Network Coverage, Reverse Auctions, Ubiquitous
Sensing, Pervasive Computing

Copyright © 2012, Luis G. Jaimes

ACKNOWLEDGEMENTS

I take this opportunity to express my sincere thanks to Dr. Miguel Labrador, for giving me this wonderful opportunity of working on this project. I am also grateful to him for his extended support and guidance throughout the course of this work, and for making my study at USF a pleasant and exciting educational experience. My sincere thanks to Dr. Sarkar and Dr. Moreno, for being in my committee and for their valuable comments and suggestions.

It takes more than words to express my thanks to my family for their constant motivation and support, without which this work would not have been possible. I thank all my friends for their continuous encouragement and support.

TABLE OF CONTENTS

LIST OF TABLES	iii
LIST OF FIGURES	iv
ABSTRACT	v
CHAPTER 1 INTRODUCTION	1
1.1 Motivation	1
1.2 Problem Statement	2
1.3 Contribution of this Thesis	3
1.4 Organization of this Document	3
CHAPTER 2 BACKGROUND AND LITERATURE REVIEW	4
2.1 Similar Concepts	4
2.1.1 Place of Participatory Sensing	5
2.2 Design Issues for Incentive Mechanisms in PS	7
2.2.1 Economically Feasible	7
2.2.2 Data Quality	7
2.2.3 Area Coverage	8
2.2.4 Fairness	9
2.2.5 Adequate Number of Participants	9
2.2.6 Adaptable to Increased Demands	10
2.2.7 Independent / Human Controlled	10
2.3 Taxonomy	11
2.3.1 Monetary Incentives	11
2.3.1.1 Monetary Static Incentives	12
2.3.1.2 Monetary Dynamic Incentives	13
2.3.2 Non-monetary Incentives	18
2.3.2.1 Collective Incentives	19
2.3.2.2 Social Incentives	20
2.3.2.3 Social Interaction Incentives	21
2.3.2.4 Intrinsic Incentives and Fun	22
2.3.3 Self-benefits	23
2.4 Performance Evaluation of the Current Systems	24

CHAPTER 3	THE GREEDY INCENTIVE ALGORITHM	26
3.1	Geometric Coverage Model	26
3.2	Relationship Between Geometric and Auction Models	27
3.3	The Greedy Budgeted Maximum Coverage Algorithm (GBMC)	27
3.4	Budget	28
3.5	Advantages of the GBMC Algorithm	30
CHAPTER 4	IMPLEMENTATION AND EXPERIMENTAL RESULTS	33
4.1	Modeling the PS System Components	33
4.2	Performance Evaluation	35
4.2.1	Experiment 1: Determining the Ideal Length of Radius R	35
4.2.2	Experiment 2: Comparing the Performance Metrics	37
4.2.3	Experiment 3: Comparing Coverage	37
CHAPTER 5	CONCLUSION AND FUTURE WORK	41
5.1	Conclusions	41
5.2	Future Research	42
5.2.1	Costs and Utility Functions	42
5.2.2	Schemes that Support Different Models of Sensing	42
5.2.3	Type of Incentive Mechanism	43
5.2.4	Working Autonomous, Intelligent Agents	43
5.2.5	Economic Models	43
5.2.5.1	Proposal for a Data Market	43
REFERENCES		46
APPENDICES		51
Appendix A	Permissions	52
Appendix B	Common Terms and Abbreviations Used in this Thesis	53

LIST OF TABLES

Table 2.1	Classification of the Type of Incentives in Terms of Design Issues	25
Table 4.1	Simulation Parameters Copyright © 2012 IEEE [1]	33
Table B.1	Common Terms and Abbreviations Used in this Thesis	53

LIST OF FIGURES

Figure 2.1	Frameworks of Human Participation	6
Figure 2.2	Taxonomy of Incentives for Participatory Sensing	12
Figure 2.3	RADP-VPC-RC Algorithm	14
Figure 2.4	Dynamic Pricing Reward Model	18
Figure 3.1	Example of Coverage per User	27
Figure 3.2	The Proposed Incentive Mechanism	31
Figure 4.1	Radius vs. Number of Active Participants Copyright © 2012 IEEE [1]	35
Figure 4.2	Radius vs. Budget Assigned per Round Copyright © 2012 IEEE [1]	36
Figure 4.3	Experiment 2 - GIA vs. RADP-VPC-RC Copyright © 2012 IEEE [1]	39
Figure 4.4	Experiment 3 - GIA vs. RADP-VPC-RC Copyright © 2012 IEEE [1]	40
Figure 5.1	Proposal for a Incentive Architecture for PS	45
Figure A.1	Permission	52

ABSTRACT

Participatory Sensing (PS) systems rely on the willingness of mobile users to participate in the collection and reporting of data using a variety of sensors either embedded or integrated in their cellular phones. Users agree to use their cellular phone resources to sense and transmit the data of interest because these data will be used to address a collective problem that otherwise would be very difficult to assess and solve. However, this new data collection paradigm has not been very successful yet mainly because of the lack of incentives for participation and privacy concerns. Without adequate incentive and privacy guaranteeing mechanisms most users will not be willing to participate. This thesis concentrates on incentive mechanisms for user participation in PS system. Although several schemes have been proposed thus far, none has used location information and imposed budget and coverage constraints, which will make the scheme more realistic and efficient. A recurrent reverse auction incentive mechanism with a greedy algorithm that selects a representative subset of the users according to their location given a fixed budget is proposed. Compared to existing mechanisms, GIA (i.e., Greedy Incentive Algorithm) improves the area covered by more than 60 percent acquiring a more representative set of samples after every round, i.e., reduces the collection of unnecessary (redundant) data, while maintaining the same number of active users in the system and spending the same budget.

CHAPTER 1

INTRODUCTION

1.1 Motivation

Studying large-scale variables in nature such as, temperature, humidity, noise or pollution requires the collection of a vast amount of data across a variety of locations. Today, this data is collected and processed by means of sophisticated, accurate, and expensive sensors located in meteorological stations. Since the high costs and the location constraint of this equipment, many cities do not have the necessary number of stations to satisfy the coverage demands. To overcome the limitations, these systems use complex statistical methods of interpolation in order to estimate the value of the variable of interest (i.e., temperature, humidity, pollution) in places located outside of the stations' scope. However, factors, such as long distance among stations and sudden changes in the environmental conditions, may compromise the accuracy and reliability of these estimates. An appealing alternative to this way of gathering environmental measurements is the involvement of citizens in the process of collecting data. Lately, this way of gathering has been made popular by the recent introduction of schemes of citizen participation such as Participatory Sensing (PS). PS leverages the sensors embedded in the citizens' mobile phones, allowing them to sense their surrounding environment. Given the number, mobility and pervasive characteristics of large populations, the resulting data collected by them turned out to be more representative and accurate than those gathered by a few meteorological stations. Additionally, the cost reduction of both implementation and maintenance as well as the ease of implementation, make PS a real alternative to the current way to sense the environment.

In spite of its multiple advantages, PS is heavily dependent on citizen participation; people are the system's core. One of the main challenges to be addressed for the successful deployment of PS

systems in real scenarios is ensuring the citizens' participation. Since citizens do not have a direct benefit from reporting data, the inclusion of incentive mechanism to encourage their participation is essential.

This thesis presents the Greedy Incentive Algorithm (GIA), a reverse auction-based incentive mechanism based on the Reverse Auction Dynamic Price with Recruitment (RADP-VPC-RC) mechanism proposed in [2], that not only includes provisions to retain the users but also algorithms to obtain the lowest cost samples that are best distributed to cover the area of interest within a given budget. This last part of GIA is achieved by the Greedy Budgeted Maximum Coverage (GBMC) algorithm [3], allowing GIA to improve coverage by more than 60 percent, acquiring a more representative set of samples after every round.

1.2 Problem Statement

This thesis addresses two fundamental challenges in the design of incentive mechanisms for PS: make the system economically feasible from an economical point of view and guarantee the area coverage. The former challenge is a necessary condition for the sustainability of any project. Here a global budget is evenly divided by the number of rounds, the procedure is presented in detail in Section 3.4. In the latter case, ensuring the area coverage of the target area is a necessary condition to obtain a representative set of sensing samples of the area of interest. After reviewing the proposed incentive mechanisms for PS up to the current time, this thesis did not find any mechanism that took the user location and a finite budget constraints into account. In order to make the system economically feasible, the incentive mechanism must consider how to maximize a finite budget in order to reward the largest possible number of participants. On the other hand, in order to acquire a set of representative sensing samples (e.g., pollution samples of the city) the samples must be taken in different places of region of interest so that the entire area can be covered by these samples. For that reason, the participants location must be considered as a criteria for data acquisition.

One approach normally used to reward participants and guarantee their participation is based on reverse auctions [2]. This approach is also used here. One of the advantages of reverse auctions is that they involve participants in a race to sell their sensing samples and the auctioneer generally

purchases the cheapest m ones, reducing the cost of maintaining the system working. On the down side, in reverse auctions users may drop out of the system, jeopardizing the system altogether, and therefore special mechanisms must be included to avoid this situation.

Although it is economically advantageous to buy the m samples with the lowest cost in every auction round (i.e., maximizing the budget), the users may all be physically located very close to each other, rendering pretty much the same information. In contrast, an incentive mechanism with location and coverage constraints will not only buy the least expensive samples but also the ones that are better distributed along the area of interest. Also, it is not practical to assume that the auctioneer will always have an unlimited budget to run the system, and therefore, including a limited budget per round makes more sense.

1.3 Contribution of this Thesis

One of main contribution of this thesis is a mechanism of incentives for PS that take into account a finite budget and participants' location. Besides, the second chapter presents a survey that summarizes the different PS incentive proposals up to the present time. This work also presents a set of design issues and a qualitative evaluation of the current incentive mechanism for PS. On the other hand, the role of PS is proposed in an emerging landscape of similar concepts involving citizen participation. Finally, a three-level taxonomy is proposed, based on the type of user motivation used for the current PS incentive mechanisms.

1.4 Organization of this Document

The rest of the thesis is organized as follows: Chapter 2 explores the place of PS in a recent landscape of frameworks that encompass it and similar tasks; presents the state of the art incentive schemes for PS; and propose a set of design issues as well as a taxonomy of current schemes. Chapter 3 discusses the algorithms' design, elaborates on the parameters, provides a functional specification, and describes the GIA algorithm components. Chapter 4 presents GIA's performance evaluation which include a set of three experiments. Finally, Chapter 5 concludes the thesis and provides directions for future research.

CHAPTER 2

BACKGROUND AND LITERATURE REVIEW

This chapter presents the state of the art incentive schemes for PS. A set of design issues are proposed as well as a taxonomy of current schemes. The rest of the chapter is organized as follows: Section 2.1 explores the roll of PS in a recent landscape of similar concepts. Section 2.2 presents a set of design issues or desired requirements of PS incentive schemes. Section 2.3 presents a taxonomy of the available incentive mechanisms for PS. Finally, Section 2.4 presents an qualitative evaluation of the different systems, and the advantages and disadvantage of each one based on the proposed design issues.

2.1 Similar Concepts

The revision of the state of the art incentives mechanism in PS has led to the discovery of a variety of concepts such as Human Computing (HC), Crowdsourcing (CDS), Social Computing (SC), Collective Intelligence (CI), and Citizen Science (CS), all of which seem to include PS and always appear linked to it. This section gives a set of brief descriptions of each of these concepts, and proposes the place of PS in this new landscape of related concepts.

Participatory Sensing is the process through which community members use the sensors embedded in their personal belongings (e.g., smart phones, bicycles, shoes, cars, etc) to measure a complex spatial phenomenon in order to discover underlying patterns. [4] Human Computation is a computing paradigm leverages the human capacity to solve some problems that computers does not handle very well. A typical example is the human capacity to solve CAPTCHAS. CAPTCHA is program that picks a random string of letters and then it renders the string into a randomly distorted image. This program generates a test which consists of asking the user to type the characters that

appear in the image. Humans can do it easily, however machines cannot. Another, term close to PS is CrowdSourcing which is a mechanism used for several companies that consists of outsourcing a task that used to be carried out by regular employees in order to reduce production costs and avoid costly payroll. [5]

Citizen Science is also a popular model of citizen participation. This participation make possible to accomplish research goals that otherwise would not be possible. Some of the most common tasks carried out for the voluntaries include collection and data analysis, testing, and some times scientific dissemination. Other models such as Social Computing emerged from the concern about social behavior. This model is often understood in two different ways. On one hand, the concept may be refereed as a way to support social behavior; this is the case of tools such as email, social networks, blogs, and in general, all the computational ways to support human interactions. On the other hand, this concept may be refereed to the different ways to support the computation carried out by collectives. Some examples include prediction markets, online auctions, and reputation systems. [6]

Finally, the concept of Collective Intelligence is related with the knowledge that emerges for the competition and interaction within groups. Although the members may not work together in a coordinate way, an agreement seems to emerge. Financial markets are a good example to illustrate this concept. In this case the stock price is not set just for an individual or in a coordinate way, instead it emerge from the independent work of many people all of them working for their own benefit. [6]

2.1.1 Place of Participatory Sensing

According to Paulos *et al.* [7, 8] the use of mobile phones as a measurement instrument for environmental sensing is just one of several methods used in CS campaigns. Therefore, it is considered PS completely included in this framework. On the other hand, PS seems to intersect HC in the sense that PS is clearly an alternative to the traditional way of sensing the environment (e.g., metereological stations) which involve machines and algorithms that lack the number and mobility of human beings. However, authors such as Quinn and Bedersonm [9]; Yuen *et al.* [10]; and Doan *et al.* [11] emphasize that the human participation in HC and CDS is directed by the computational

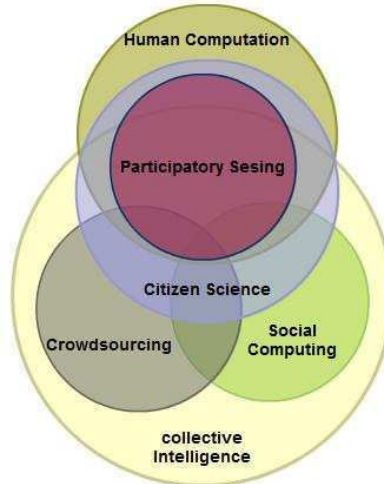


Figure 2.1: Frameworks of Human Participation

system or process. From this point of view, when the PS task is carried out by a spontaneous group of people with no clear guidance (e.g., using social networks) it can not be considered as part of HC nor CDS. Besides, part of PS appears to be in the intersection of HC and CDS. On the one hand, today the sensing process in PS may be carried out by machines in opposition to HC, but on the other hand, the same task may be done by training employees (e.g., in opposition to CDS). In both cases PS is an alternative for replacing either one, the machine mode or the employee mode, which results in a trade off between quality and costs.

Figure 5.1 illustrates the place of PS in this landscape of related concepts; collective intelligence is a superset of social computing and crowdsourcing, because both are defined in terms of social behavior. Whereas human computation replaces computers with humans, crowdsourcing replaces traditional human workers with members of the public. The key distinctions between collective intelligence and human computation are the same as with crowdsourcing (i.e., the system coordinates human participation). However, there is an additional distinction that collective intelligence applies only when the process depends on a group of participants. According to Quinn and Bederson [9] given that the definition of PS does not say anything about groups, there could be an PS system of a single worker in isolation. This is the reason that part of PS is outside of CI.

2.2 Design Issues for Incentive Mechanisms in PS

The following characteristics were identified as necessary conditions for incentive mechanisms for PS: Economic feasibility, data quality, area coverage, fairness, adequate number of participants, adaptable to increased demands, and finally, works independent or dependent of the human control.

2.2.1 Economically Feasible

For any project, budget is a critical factor. Activities like PS need a critical mass of people to maintain both reliability and coverage. However, budget concerns constrain these goals. Hence, a key design issue is to maintain an equilibrium between reliability and coverage. A common approach to address this problem is to estimate the *reservation wage*, which is defined as the lowest wage rate at which a worker would be willing to accept a particular type of job. That means that a job offer involving the same type of work and the same working conditions, but at a lower wage rate, would be rejected by the worker. However, according to Addison *et al.* [12] an individual's reservation wage may change over time depending on a variety of factors. Taking into account the dynamic nature of this indicator, the introduction of additional components such as entertainment and competitiveness may contribute to lower the reservation wage. Section 2.3 introduces some dynamic monetary incentives for PS that combine monetary with intrinsic motivations to make economically feasible the system deployment as well as non monetary mechanisms that use other types of motivations to encourage user participation.

2.2.2 Data Quality

Information is the core of a PS system, and its incentive mechanism must ensure that potential consumers receive a higher Quality of Information (QoI) according to their requirements. A natural question is, how can a PS incentive mechanism encourage the collection of quality data? One of the most common approaches utilized to address this problem is the use of reputation schemes. Generally, users' reputations may be computed from their past performances, the assessments of their peers, or by a combination of those two. Yang *et al.* [13] proposes the use of this criteria to rank participants in four levels that go from very trustworthy in the top 25% to very untrustworthy

in the bottom 75% - 100%. On the other hand, Huang *et al.* [14] uses the Gompertz function to rate the trustability level of the data provided by participants. In both cases, these reliability indicators can be linked to credit incentives in order to encourage the user movement from low to top quality levels. Once computed, this indicator (e.g., trustability index) may be used to predict quality. Pham *et al.* [15] proposes the use of those indicators as a criteria for data acquisition. Pham *et al.* [15] models the selection of participants as a multi-objective optimization problem, where quality is one of the objectives and the outcome is the best trade off between quality and *cost*. Reddy *et al.* [16] propose a feedback reinforcement method to assess metrics such as: timeliness capture, relevancy, coverage, and responsiveness. The feedback component continuously reinforces the improvement of these metrics by the use of monetary and social-base incentives. Finally, Liu *et al.* [17] propose an *on demand* method for data acquisition, in which each query is associated with a required quality level. User's credit reward is linked to the QoI satisfaction index of the data customer. In this way, depending of the query quality level requirement, it is directed to the group that may provide this quality level.

2.2.3 Area Coverage

Another important goal of any PS incentive mechanism is always selecting the best subset of users in order to estimate a complex spatial phenomenon (e.g., temperature, pollution, noise levels). To address this problem Reddy *et al.* [18] propose the use of mobility profiles as part of the participant's selection criteria in the recruitment process. Using a similar approach, Falaki *et al.* [19] and Shilton *et al.* [20] propose increasing the participant's demographic diversity and social network affiliation. They propose to leverage the mobility patterns of different groups to increase the sensing coverage area, and also envision tackling problems such as language barriers by the use intrinsic motivations. Kuznetsov and Paulos [21] propose the involvement of stakeholders such as students, parents, bicyclists, and homeless people. They say that each of these groups is disseminated on different public spaces. On the other hand, Lee and Hon [2] propose the use of Virtual Participation Credit (VPC), a mechanism for sensing samples acquisition that uses a principle of locality that assumes people with a similar reward expectation are clustered geographically.

2.2.4 Fairness

Fairness is a key factor in accomplishing goals such as retention, coverage, and economic feasibility. Usually this concept is understood as the equal opportunity of participation. However, in the case of dynamic pricing models based on reverse auctions, fairness is understood as a scheme in which a user whose bid is cheaper has a higher chance to be selected than a user whose bids are more expensive. In this case (incentives models based on reverse auctions), this feature (fairness) may lead to a geographical coverage imbalance as well as user drop out. In the former case, because the participants with a lower reservation wage may be located in a particular region of the target area, and in the latter case, because the samples will always be acquired from those with a lower reservation wage causing that, after several periods of waiting, those with a higher one drop out because they do not have the opportunity to participate. Some authors, such as Lee and Hon [2] address this problem by the use of mechanism such as VPC and Recruitment (RC) which avoid always acquiring samples from the same users (e.g., the cheapest ones). On the other hand, Pham *et al.* [15] tackle this situation using a mechanism that not only considers the sample's price, but also its quality. Finally, Jaimes *et al.* [1] approach the same problem by means of a greedy algorithm that not only considers the sample's price, but also their location, which allows that the users' location as well as their mobility patterns, increases or decreases their chances to be chosen.

2.2.5 Adequate Number of Participants

Because the number of participants and their mobility patterns, PS seems to be an appealing alternative to traditional sensing. Factors such as sensing frequency, target area size, type of phenomenon to be analyzed, and sensing demands can be used to determine the necessary number of participants for a specific sensing request. However it is always necessary to have an additional number of participants, given the uncertainty in terms of size, location and accuracy of next request of sensing (e.g., sensing on demand) and the natural dropping out of participants (e.g., high frequency sensing). In some monetary based schemes, in which a budget constraint determines the number of selected users, the starvation of participation may cause dropping out of them. This is the case of [2, 15] in which the users must send the sensing data in advance, and then the system

decides based on factors like quality, whether to acquire them or not. In this case, the extra burden of collecting and transmitting data without any reward may cause the user to drop out after a few rounds. The inclusion of reputation schemes to predict data quality such those reviewed in section 2.2.2, may avoid this drawback. In this way, all the subscribers of a Phone Network Provider (PNP) are potential participants of the PS system, and their participation depends on whether their reward expectations are satisfied.

2.2.6 Adaptable to Increased Demands

Scalability can be defined in terms of how well a solution to problem will work when the size of the problem increases. From this point of view, the incentive systems must have the ability to maintain performance, usefulness, and usability regardless of expansion and concentration of participants from local to a more distributed pattern. This expansion implies not only both an increasing number of participants and the inclusion of new target areas, but also the inclusion of new types of participants in terms of their intrinsic interests and expected rewards. It may be necessary to use methods that dynamically adapt to different types of population according to their motivations. In terms of number of nodes and area coverage, PNP are able to scale to millions of nodes, which implies almost a total coverage. However, the ability to attract and maintain a wide range of users based on their different types of motivations, has not yet been sufficiently developed. Approaches such as Micro-Blog [22] combines the social incentives supplied by the use of social networks with a type of incentive called “give and take” which encourages respond queries when they are generated by strangers.

2.2.7 Independent / Human Controlled

An equally important design issue has to do with the ability of the system to work autonomously in the absence of human intervention, i.e., continue encouraging the user participation in scenarios in which the user acts as passive carrier of the sensors. Lane *et al.* [23] and Kapadia *et al.* [24] classified the grade of user’s participation as opportunistic and participatory. In the former case, users volunteer their cellphones to transparently sense the environment while they continue with

their daily lives without being interrupted. They may not be aware of when or how their cellphones are remotely tasked to collect and report data. In the latter case, users are actively involved in the sensing process (e.g., taking a picture, manually activating a sensor, manually accepting or rejecting a request for sensing). The emergence of hybrid systems such as [25, 26] that combine these two approaches, leads to incorporating new characteristics like the inclusion of agents that automatically negotiate on the user's behalf. Sun and Nakata [27] propose a an agent-based architecture for PS to overcome the lack of user's active participation. This architecture includes four types of agents: user agent, local agent, data verification agent, event monitoring agent, and central agent. User agents provide raw data, which are processed by the local and verification agents, which in turn send it to the database. The event monitoring agent analyzes and detects database anomalies and potential hazards. Finally, the central agent is responsible for receiving and processing every query. A detailed overview of the role that systems of intelligent agents play in the design of dynamic incentive mechanisms is presented by Parkes *et al.* [28].

2.3 Taxonomy

PS incentives can be classified in different ways. One way is based on the level of user participation (i.e., Opportunistic or Participatory Systems). Since there are many incentives that apply to both of them, in this thesis, the classification is based on the types of stimuli that encourage user participation. In Figure 2.2, this classification approach initially identifies two large branches: Monetary and non Monetary incentives. The former is classified as either Static or Dynamic. The latter are divided into the following categories: Collective Motives, Social Reward, and Intrinsic Motives-based incentives. Finally, the Social Reward-based incentives are further divided as Interaction, and Self-Interest-based incentives.

2.3.1 Monetary Incentives

When the sensing activity does not bring any immediate benefit for users, monetary rewards are powerful motivators. They encourage participants to perform tasks that otherwise they would not perform voluntarily. The way the reward is determined leads to two types of approaches: Static

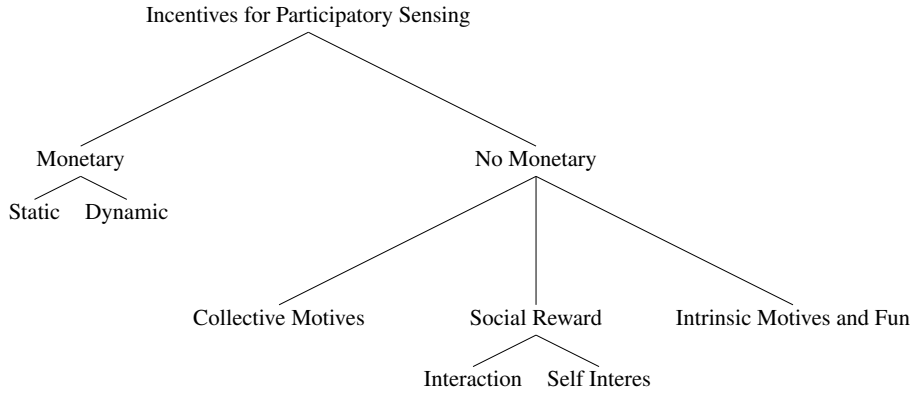


Figure 2.2: Taxonomy of Incentives for Participatory Sensing

and Dynamic. In the former, the amount to pay to each participant is estimated in advance and the amount may differ for each participant. In the latter, the payment depends on the reservation wage of each participant. A common approach is to combine monetary rewards with other types of incentives (e.g., Intrinsic Motives, Social-based incentives, etc) in order to decrease the user's reservation wage.

2.3.1.1 Monetary Static Incentives

Static types of monetary incentives include the first four of the five types of micro-payments presented by Reddy *et al.* [29]: Macro, Low, Medium, High, and Complete micro-payments. Reddy's findings showed that the way to estimate this payment affects the participation level and data quality. In terms of quality and coverage, the most successful methods reported in this study were medium and high micro-payments, and in terms of the number of submitted samples, the best method was the dynamic one (i.e., Complete micro-payment). However, this type of macro-payment also caused the highest level of users dropping out.

Musthag *et al.* [30] study confirmed these results using three types of micro-payments: Uniform, Variable, and Hidden. In the Uniform scheme, a fixed amount was rewarded for completing a micro-task; in the Variable scheme, random amounts were assigned based on a prior distribution; and in the Hidden scheme, rewards again varied, but the amount was not revealed until after the micro-task was completed. Their findings showed that in all the cases they retained around 62% of the

participants, the percent of completion rate was slightly higher in the variable approaches with 86.92% followed by 86.26% and 81.68%. On the other hand, the *Variable* approach obtained the highest level of positive participants perception followed by the *Uniform* and the *Hidden* one. Their conclusion point out the superiority strategy of use Variant over Uniform, and the poor performance of Hidden. Finally, they result confirm the findings of Reddy *et al.* [29] in the sense that the data quality is influenced by the incentive strategy.

2.3.1.2 Monetary Dynamic Incentives

Lee and Hon [2] present Reverse Auction Dynamic Price with Virtual Participation Credit and Recruitment (RADP-VPC-RC) an incentive mechanism inspired in micro-economic theory and motivated by the following scenario. There are n participants spread through the area of interest. They are willing to sell their sensing data to an auctioneer who in turn wants to acquire the m less expensive samples, within a time window (i.e., rounds). The same process is repeated for k different rounds and is carried out through a reverse auction. The first component (RADP) is a recurrent reverse auction that dynamically reduces the reservation wage of each participant. Figure 2.3 shows the RADP-VPC-RC process overview. Once the auctioneer buys the m cheaper samples, the winners (i.e., those who managed to sell their samples) increase their bid's price by 10% expecting to increase their profits and losers decrease it by 20% in the hope of winning in the next round. In order to model the drop out, losers (i.e., those who did not manage to sell their samples) evaluate their ROI value (i.e., which is based on their reservation wage and sum of the cost that the user incurs in collecting the data) which is represented by equation 2.1. If it is below a certain threshold, then they withdraw from the system, otherwise they continue participating in the next round.

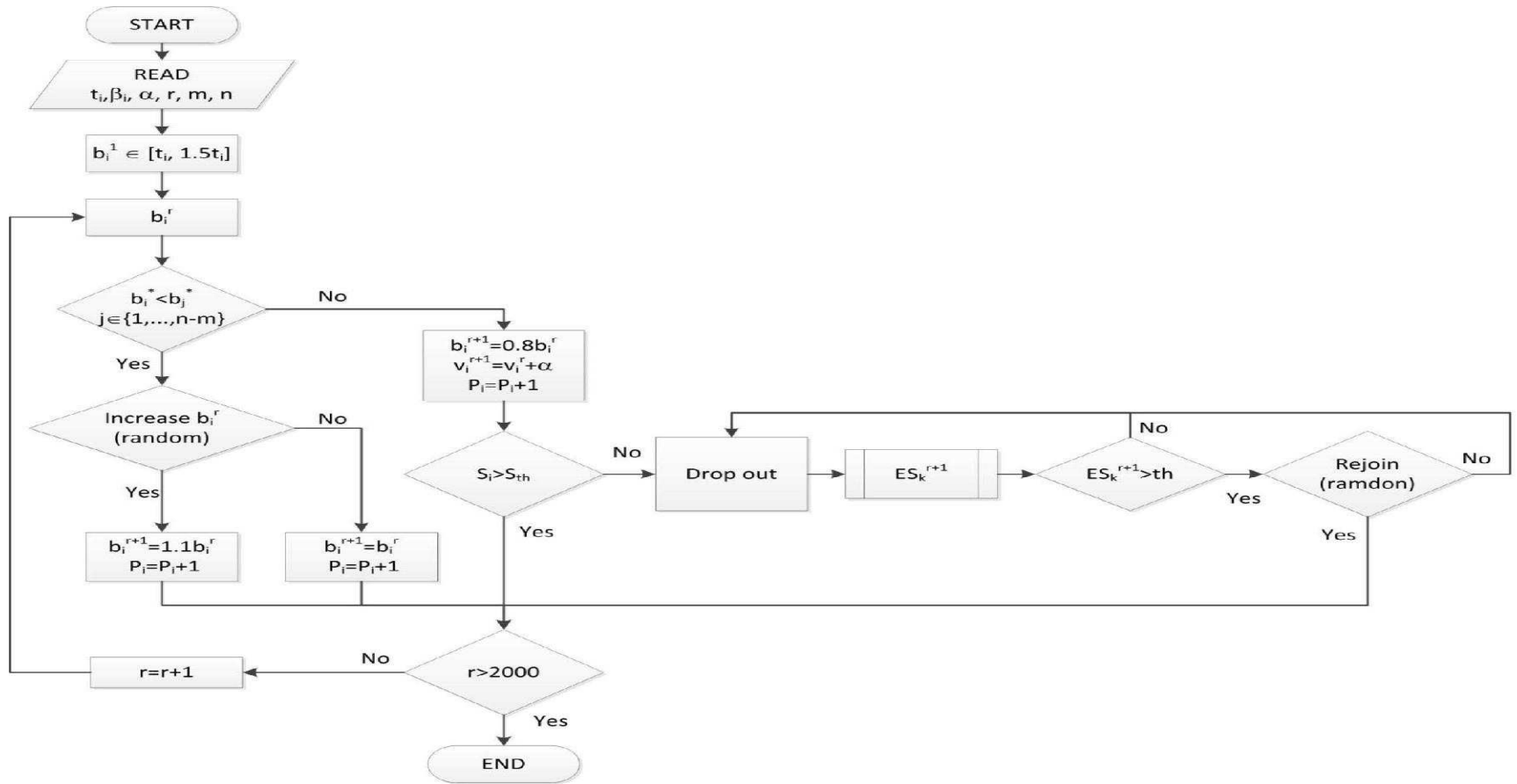


Figure 2.3: RADP-VPC-RC Algorithm

$$S_i^r = \frac{e_i^r + \beta_i}{p_i^r \cdot t_i + \beta_i} \quad (2.1)$$

Here, e_i^r corresponds to the earned reward by users i until round r , $p_i^r \cdot t_i$ corresponds to the minimum reward, with p_i^r as the number of participation instances of i up to the current auction round r , t_i as the user's true valuation (reservation wage), and β_i as the tolerance period. In order to increase the level of participation, users who lost in the previous round $r - 1$ and continue participating, grants a VPC $v_i^r = v_i^{r-1} + \alpha$, where α corresponds to the amount of VPC. Hence, losers increase their chances to win by using their virtual bid's price $b_i^{r*} = b_i^r - v_i^r$ rather than the real bid's price. Additionally, the RC mechanism allows the auctioneer to communicate the maximum price paid φ_k to a winner in the last round to the users who dropped out in previous rounds. This information allows them re-evaluate their *ROI* and potentially return to the system in the next auction round. This expected *ROI* is evaluated as follows:

$$S_k^{r+1} = \frac{e_k^r + \varphi_k + \beta_k}{(p_k^r + 1) \cdot t_k + \beta_k} \quad (2.2)$$

This economic model avoids two typical problems of reverse auction: dropping out and cost explosion, making the project economically feasible by means of the combination of RADP, VPC, and RC. However, others problems such as: coverage and data quality are not totally addressed. The former, is weakly tackled by the use of VPC and by a proposal of applying the RADP-VPC-RC simultaneously in different regions, and the latter is not addressed at all. However, other proposals such as Pham *et al.* [15] and Jaimes *et al.* [1] based on RADP, address those problems by using multi-objective Knapsack and Budgeted Maximum Coverage approaches respectively.

Inspired by a similar scenario as above, Pham *et al.* [15] addresses the problem of selecting the samples that will be acquired in each round, minimizing the ratio between the total price paid for the acquired samples and its total quality (e.i., it is assumed that the network provider receives the samples, its price, and they determine its quality before acquiring them). This NP-problem is modeled as an instance of a multi-objective Knapsack, where the constrain m is the number of samples to be acquired, and objectives both maximize the quality of the acquired data and minimize the costs.

They solve this multi-objective optimization problem by the use of techniques of evolutionary algorithm. This proposal includes three steps in which the Univariate Model Distribution Algorithm (UDMA) [31] and Third Evolution step Generalized Differential Evolution (EGDE3) [32] algorithms are used to (1) select the initial population by a pair-wise tournament, (2) Use UMDA to create the next generation, and (3) transform individuals from infeasible to feasible states. This evolutionary process is repeated and the resulted individuals are the input for the next generation. Their simulation results show that the proposed scheme acquires samples of high quality by 10% less money than RADP and up to 15% less than Genetic Algorithms (GA) scheme. Besides, this hybrid scheme solves the Knapsack problem three times faster than GA.

Another, representative approach is presented by Liu *et al.* [17]. This work presents a network management framework to tackle three fundamental problems: maintaining energy supply, supporting quality information (QoI), and incentivizing the user participation so that the PNP obtains maximum revenue. Even though the solution of the three problems are related each other, the first two are outside the scope of this thesis, so here, is only dealing with the latter. The proposed framework is motivated for the following scenario. There are a set P of users subscribed to a PNP, and a user (querier) that additionally owns a subscription to a PS plan offered by the company. Any time the querier asks the PNP for contextual information about a site of interest k , the PNP in turn re-directs the query to a set of N users within the proximity of k , with $N \subseteq P$. Once the query is received, the users who decided to participate sending their sensing data, are rewarded based in the PNP's profits. In this dynamic pricing scheme, unlike the RADP-VPC-RC, the user's reward expectation may be unknown in advance and must be predicted. The exponential smoothing method is used to predict the current expectation reward of each user. This method uses the most recent payment history to compute the current expected reward and the degree of satisfaction expressed by the users who received this payments to adjust this value. Equation 2.3 shows the expected reward for user i , who answers a query q with r attributes required.

$$\hat{\varphi}_q^{i,r} = (1 - \mu)\hat{\varphi}_{q-1}^{i,r} + \varphi_q^{i,r}, \forall i \in N, q \in Q, \quad (2.3)$$

where $\widehat{\varphi}_{q-1}^{i,r}$, $\varphi_{q-1}^{i,r}$, and $\mu \in (0, 1)$, correspond to the predicted and actual reward requirements for the query $q - 1$ respectively and $\mu \in (0, 1)$ to a weight factor. On the other hand, the system compute the exact participant's reward, minimizing the sum of mean square error of the predicted value computed in Equation 2.3 as follow:

$$\omega_q^* = \arg \min \frac{1}{N} \sum_{i \in N} \{\widehat{\varphi}_q^{i,r} - \omega_q(\widehat{c}_q^i)\}^2 \text{ subject to: } \left\| \frac{\omega_q - \omega_{q-1}}{\omega_{q-1}} \right\| \leq \delta \quad (2.4)$$

where δ represents an upper bound in terms of the percent of adaptation from PNP and \widehat{c}_q^i is the type of information with which the user i responds to query q . Figure 2.4 sketch the final reward computation process.

This framework addresses issues such as project's economic feasibility, data's quality assurance, and requirements in terms of the number of participants. On the other hand, each query is associated with r levels of QoI attributes. Once the PS system receive the answers to the query q , it fusion all of them in order to offer a higher level QoI answer to the querier. Besides, quality indexes such as μ_q^r and μ_q^q (i.e., required and attained degree of event understanding) are used to measure the QoI satisfaction index. These ways to measure quality satisfaction allow easily identify quality problem in the systems. On the other hand, the retaining is addressed by a combination of dynamic credit reward and a reduction of impact of the participants power consumption. Other issues like are coverage are not totally address, but can be inferred that it depends of PNP coverage service.

Finally, Xie *et al.* [33] propose an incentive mechanism motivated by an intermittent connection environment like Delay-tolerant networks. Here, each query for contextual information is sending to a group N with n members, within the proximity of querier interest site. Each query is associated with a tuple (r, s_r, m) which include the type of message r ; a specific sink s_r which is the only node authorized to receive data of type r from nodes in N and deliver these data back to the querier; and a message identification sequence number m . Given the nature of the network (i.e., intermittent and low level connection) it is possible that the mobile node that obtained the information does not have the opportunity to meet with the corresponding sink, and instead, have occasional encounters with other nodes. Besides, the sink only rewards the node from which it get the data (i.e., even, if this node was not the data generator), only if a copy of the data has not been acquired so far. In

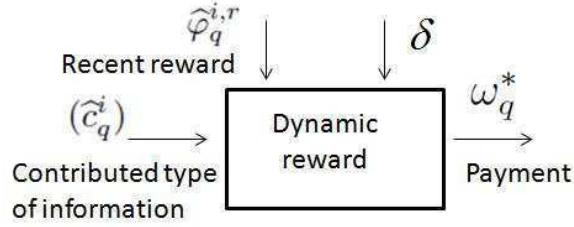


Figure 2.4: Dynamic Pricing Reward Model

this context, the proposed scheme incentivize a collaborative data trade between nodes, so that, the information finally being reached by the sink, and it in turns can delivered it in a safe way to the querier. This data trade is expected under the assumption that nodes are selfish and they want to increase its profits. Therefore, they want to trade copies of different types of data, so that, when a node meets a sink, this node has high chances to be rewarded.

Let $P_i(r)$ denoted the probability that node i has an encounter with the sink. Let $A_i^m(r)$ be the message appraisal, which is probability that other nodes, but i has not delivered any copy of the message m to the sink that support r type of messages. Once two nodes i and j have contact, they start the bargain process determining which information will be traded. This process is modeled as a two-person cooperative game, and a Nash equilibrium is applied to reach the optimal solution. Equations 2.5 shows the message appraisals values for nodes i and j before and after i sends a message with sequential number m of type r to j . At this moment both node have a copy of m .

$$\begin{cases} A_i^m(r) = [A_i^m(r)] (1 - P_j(r)) & \text{sender} \\ A_j^m(r) = [A_j^m(r)] (1 - P_i(r)) & \text{receiver,} \end{cases} \quad (2.5)$$

Finally, the i node expected credit reward is given by equation 2.6

$$R_i^m = A_i^m(r) \times P_i(r) \quad (2.6)$$

2.3.2 Non-monetary Incentives

These type of incentives have been the subject of intense study and research. These type of incentives allow the voluntary participation of citizens in research projects that otherwise would not

be possible to carry out because its high costs of execution. Agencies like NFS strongly encourage the use of these type of incentives in activities such as: gather information, data collection, analyze samples, and analyze data, through framework such as CS [34] and PS. Some well-known examples of this type of projects include Galaxy Zoo [35] (i.e., a project that involve the participation of citizens in the classification of galaxies), and eBird *et al.* [36] (i.e., a project headed by Cornell Lab of Ornithology). These types of incentives have reached its climax with the advent of social networks, the mass use of blogs, the new smartphones' communication capacities, and in general, technologies that support the interaction among members of networks and communities. Part of the proposed classification of Non-Monetary Rewards-based incentives has been influenced by the theoretical work in social movement participation of Klanermans and Stekelenburg [37], and Simons *et al.* [38] as well as the research of Nov *et al.* [39] who after surveyed 4376 volunteers, found that collective, intrinsic and social-reward incentives were most significant factors of participation. In the context of PS this thesis has identified two subsets of social rewards incentives: Interaction and Self Interest. In the former case, rewards comes from the interaction with other members of the network, among them are, recognition, new friends, reputation, and felt part of a community. In the Latter, the reward comes from the use of the resources created by the community for personal benefit.

2.3.2.1 Collective Incentives

This type of incentive encourages the attainment of a common good. This incentive was identified as primary motivation in PS, because ultimately this is the philosophy of PS, work together for a common good (i.e., a better environment, a better community, and better planet). Once reached the objective, everybody will benefit from the goal achievement, no matter whether they have participated or not. For instance a public PS campaign to measure pollution levels, will improve the wellness of all the members of the community no matter whether they have or have not participated personally. According with Simon *et al.* [38] the collective good may be insufficient as a motivating force, because people can hope for a "free ride." In addition to the collective good, potential free riders may also need additional incentives, which may have the form of social rewards, economic

rewards, intrinsic motivation, and fun an entertainment. To this category belongs all the PS systems which do not mention any type of incentive, because it is understood by default that the people will participate for a common good.

Dantec *et al.* [40, 41] present a system oriented to the care in challenging environments. The system leverage the advantages of the cellphones to provide care and shelter for homeless mothers. Providers, staff, and people in need leverage the infrastructure to improve the logistic, that include the grocery bargain hunting by the providers and allocation the resources by the staff.

Mendez *et al.* [42] present P-Sense a PS system for air pollution monitoring. The system include sensors to measure levels carbon dioxide, combustible gas, carbon monoxide, and air quality as well as temperature and relative humidity. The benefits in terms of wellness and conservation of the environment from this system are part of the collective interest and represent a common good.

Kanjo [43] presents NoiseSPY a PS platform for urban noise monitoring. This system measure noise levels in order to generate map of sound levels that may be used for citizens and urban campaigns to improve the noise contamination.

2.3.2.2 Social Incentives

Social motivation play a powerful roll in user participation. Having the opportunity to interact with other members of the community, being aware about what they are doing, gaining reputation, taking self benefit from content create by the community, and use the network as a self presentation tool [44] are just some of the reasons since millions of people join each year to social networks. In a survey about the motivating potential of the features for increasing participation in PS [26]. Answers were ranked from the least motivating (1) to the most motivating (10); features such as being able to publicly comment on other participants observations (6.9 ± 2.8), and being notified of other participants making observations on items that you had first observed (6.7 ± 3.2), indicating that social networking aspects were highly desired. This results, along with other studies, suggests that interaction among members of the community through social networks such as: Facebook [45], Twitter [46], and Flickr [47, 48] is a strong way to incite people to take part in schemes of PS.

2.3.2.3 Social Interaction Incentives

The use of technologies such as social networks, blogs, SMS, email, chat, VOIP have become part of daily routines of millions of people around the world, as well as the primary way for social interaction. PS systems like CenceMe [49] use the concept of "sensing presence" which allow participants of social networks share his/her status in terms of activity, disposition, habits, and surrounding (e.g., temperature, ozone levels, noise) in a non-verbal way, which leads to new levels of connections. This system leverage the incentive characteristics of participation in social networks, the number of participants can growing at a similar rate as the members of social networks, coverage and scalability are guarantee for the same mechanism. The system is feasible from an economical point of view, and finally the system works independently and under human control.

Another PS system that use this type of incentive is NoiseTube [50], a monitor noise pollution application that allows participants measuring and sharing the levels of noise around them. This system apply the concept of social translucence [51] as a mechanism to support social interaction. This concept based on visibility, awareness, and accountability is used by NoiseTube to increase the level of participation via social stimulation. Participants can exhibit their contribution via Elog, an environmental log which allows participants see maps of noise, show its contribution, and receive feedback. Furthermore, the system can be embed as a web widget into web pages and social networks, allowing participants to be aware of his/her own and others participation, receive comments, and make the information public with which in turn they get self promotion.

Micro-Blog [22] propose to handle incentives by means of a combination of social and explicit mechanism. The former is address by the use of social networks, which makes possible the participation by the interaction of the social groups (e.g., friends in Facebook). However, in order avoid confining the participation in social cluster, they propose the use of a pool of n free query credits for each user (i.e., explicit mechanism). Each time a user make a query or reply to one, their credits are decreased or increase respectively.

PIER [52] also use a model of incentives that include social interaction through social networks, a careful manage of privacy, and control of the data shared. Users may compare the levels of pollution exposure via facebook application module, and make and receive commentaries from

other users. Authors envision include mechanisms of reputation to improve the reliability in the share information.

One advantage of these incentive mechanism is that the PS system inherits all the advantages in terms of recruitment and participation of the philosophy of social networks. On the other hand, from the economic point of view, is economically feasible maintain the user participation. However, none of the reviewed proposals mention how to cover the PS system operation costs. Followers of these type of incentive approaches envision that PS operation costs will be charged to PNP and it in turns will offer PS's capacities as part of the service package.

2.3.2.4 Intrinsic Incentives and Fun

Authors like Wiggins *et al.* [36, 34, 53] have demonstrated how users can be involved in active participation projects under specific conditions. Intrinsic Motives-based incentives are aimed at groups of people with specific interests. Typical examples of these type of incentive include games and activities such as: gardening, and bird watching among others, where the majority of participants volunteered due to their inherent interest. An example of the application of this type of incentives is eBirding [54]. This project launched by the Cornell Lab of Ornithology, in partnership with the Audubon Society, is focusing on a specific target group called birders, who range from sporadic to those whom observing birds becomes a life style. This project was initially based on a science-centric model of participation (i.e., Collective incentives in this taxonomy), but shifted to a *birder-centric* model when contributions plateaued in 2005. This new intrinsic motivation model (i.e., Intrinsic) turned out in major changes in the system, that included different features specific to the birder interest. After five years, contributions went from a few thousand to around two million observations per month [39]. This example suggests that intrinsic motivations may increase significant the level of participation when the right intrinsic incentive is used for the right target population, and the application is careful design with this purpose in mind.

On the other hand, intrinsic motivation like gaming, play an important roll in the recruitment and retention of participants. Every day, people around the world spend millions of hours playing computer games. A recent report of the International Game Developer Association (Igda) [55] reveals

that over 200 millions people, i.e., over 25% of all Internet users, play online games every week. Authors such as Von Ahn, Law, Dabbish, and Hacker [56, 57, 58, 59, 60, 61, 62, 63], [64] have pointed out about the potential of a type of games called *casual games* in which the development of tasks are part of a game. One of first works in this field, was the ESP Game [63], this is an online game that links anonymous online individuals and has them independently tag a random online image. Once they separately suggest an identical tag, they receive points for how quickly they did so. Thus, the system ends up with independently confirmed (i.e., validated) semantic information about the photo. This game with potential applications in language translation, Monitoring of security cameras, Improving Web search, text summarization according with Von [64], was licensed by Google in 2006, who made its own version called the Google Image Labeler [65]. According with Von [62] the games design with a specific purpose “*Human Computation Games*” must be constrained by the computational problem to solve; which involves a trade off between between fun and practicality.

One of recent applications of this type of incentives to PS is presented by Han *et al.* [26]. In order to measure the impact of use intrinsic incentives like gaming in PS projects, the CENS lab of UCLA chose the BudBurst project [66] as a testing platform. BudBurst is an environmental CS national project that engage citizens in the collection of ecological data. CENS team created a playing infrastructure which allow to participants tag location information about the observed plants; visualization features like maps and display of gain points; and Floracaching [67] a variant of the well-know game geocaching. A post-experiment survey asked to participants to rank from 1 to 10 in increasing order, the experiment’s most motivating elements, game was highly ranked (7.9 ± 1.5), followed for social interaction (6.3 ± 3.4). This findings show the potential of the incentives schemes based on gaming for engage user participation.

2.3.3 Self-benefits

Finally, the type of users’ incentive mechanisms based on the direct benefit obtained by for their participation were reviewed. In this type of scheme users receive immediate gratification for their participation, benefiting from content created by other participants [68]. This is the case of initia-

tives that require sharing of knowledge or experiences. By adding your own facet of information to the project, it becomes framed on a greater scale. For instance, users may submit the local price of milk, lettuce, beer, gas, etc or just take a picture of its bar code and receive an instant feedback about the price of same products in near stores. In other words, while users are making a query by sending a bar code picture, they are populating the database. This is the case of LiveCompare [69] that allows for grocery bargain hunting through PS. Other application are PetrolWatch [70] and and MobiShop [68] systems based on PS that leverage the ubiquity of smartphones to share consumer gas pricing information.

These PS system of specific purpose has an special target population, and it is a challenge incorporation this types of incentive into more general application. They have a great potential, in part because users is paying for each query with their participation, In fact, due to the utility in terms of user's money savings, they have great commercial potential.

2.4 Performance Evaluation of the Current Systems

The qualitative comparison of the incentive systems has been hindered by the fact that each system works over different scenarios, use different approaches, and works for a specific sensing model. Some of the systems correspond to operational implementations, other are pilot experiments, and some other are theoretical models tested using simulations. However, the evaluation is based on the results of the experiment, simulation, or a proposed approach. For example, those systems that use social networks, it follows that the network provides coverage. In other cases like [30] the N/A on a field of Table 2.1 indicates that this metric is meaningless in that context. Finally, The absence of X on the a field of the same table denotes that the design issue has not even been taken into account or that simulation results show a poor performance.

Table 2.1: Classification of the Type of Incentives in Terms of Design Issues

Type of Incentive			Design Issues							
Type of Incentive		Incentive Mechanism	EcoFeasi	Quality	Coverage	Fairness	Num of users	Scalability	Control	
Monetary	Static	Macro-Micro -Payments [29]	X		X	X				
		High Micro-Payments [29]	X	X	X	X				
		Medium Micro-Payments [29]	X	X	X	X				
		Low Micro-Payments [29]	X	X	X	X				
		Uniform Micro-Payments [30]	X	N/A	X	X				
		Variable Micro-Payments [30]	X	N/A	X	X				
		Hidden Micro-Payments [30]	X	N/A	X	X				
	Dynamic	RADPC-VPC-RC [2]				X	X	X		
		Dynamic Price [15]	X	X	X	X	X			
		Manag for Context-Aware [17]	X		X	X	X			
		Bargain-based [33]		X						
		Complete Micro-Payments [29]	X	X	X					
No Monetary	Social Rew	Interaction	CenceMe [49]		X	X	X	X	X	
			NoiseTube [50]		X	X	X	X		
			Micro-Blog [22]	X	X	X	X	X		
			PIER [52]	X	X	X	X	X	X	
	Self-Inter	LiveCompare [69]			X	X	X	X		
		PetrolWatch [70]			X	X	X	X		
		MobiShop [71]			X	X	X	X		
	Collective	CRM [40]			X	X	X	X		
		P-Sense [42]								
NoiseSPY [43]				X	X		X	X		
intrinsic motives	eBirding [54]			X	X	X	X			
	Floracaching [26]	X	X	X	X	X	X			

CHAPTER 3

THE GREEDY INCENTIVE ALGORITHM

This chapter discusses the algorithms' design, elaborates on the parameters, provides a functional specification, and describes the GIA algorithm with the help of a sample model.

RADP-VPC-RC and the GBMC algorithms are important components of GIA . While RADP-VPC-RC tackles the problem of cost explosion and avoids users from dropping out of the system, GBMC considers the location of the users, the coverage, and the budget constraints of the auctioneer. With these two algorithms together, GIA buys samples within the spatial and temporal coverage needed under a fixed budget. Further, GIA does not assume static users and rather includes additional mechanisms to encourage user mobility as a mean to extend the geographical coverage even further.

3.1 Geometric Coverage Model

In order to address the coverage problem, GIA uses the following geometric disk model

$$f(d(u_i, u_j)) = \begin{cases} 1 & \text{If } d(u_i, u_j) \leq R \\ 0 & \text{Otherwise} \end{cases} \quad (3.1)$$

where $d(u_i, u_j)$ is the Euclidean distance between a sensor u_i and another sensor u_j , and $R > 0$ is a constant that defines the area of coverage of each sensor. Indeed, this function defines a disk centered at sensor u_i with radius R . All sensors within such a disk have a coverage measure of 1 and are said to be *covered* by sensor u_i . On the other hand, all sensors outside such a disk have a coverage measure of 0 and are said to be not covered by this sensor.

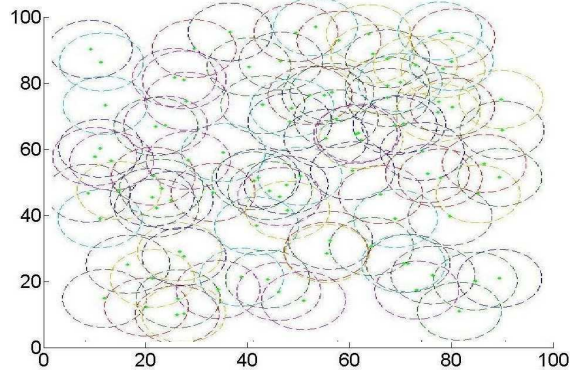


Figure 3.1: Example of Coverage per User

3.2 Relationship Between Geometric and Auction Models

One of main characteristics of the GIA for sample acquisition consists in the modification of the RADP-VPC-RC auction scheme. This auction mechanism is meant to buy k units in increasing order of cost. In this scheme, the geometric coverage model described before was used to carry out this process. Here, every sensor is drawn as a disk of radius R centered at its location, as shown in Figure 3.1. Of course, each sensor is covered by at least the disk centered at itself. The natural question here is: Does there exist a sub-collection of disks whose union cover all the points? If the elements covered by $disk_i$ are assigned to the set S_i for $i = 1, \dots, n$, where n is the number of disks, the previous problem becomes the well-known *NP-Complete Set Cover Problem*.

3.3 The Greedy Budgeted Maximum Coverage Algorithm (GBMC)

The heart of the incentive mechanism is the *Greedy Budgeted Maximum Coverage (GBMC)* algorithm, which is executed in every round of the auction to decide which samples to buy. The algorithm, which is detailed in pseudocode in Algorithm 1, finds a $S' \subseteq S$ that covers the greatest possible area covered by S within the budget constraint.

Let assign the elements covered by each $disk_i$ to each set S_i , associate the weight w_i of S_i as its cardinality, and its cost c_i as the cost of the sample provided by the sensor u_i located at the center of the $disk_i$. Let $G \subseteq S$ be a collection of sets, $w(G)$ and $c(G)$ denote the total number of elements covered by G and the total cost of the sets in G , respectively. Let $W'_i, i = 1, \dots, n$, denote the

total number of elements covered by set S_i but not covered by any set in G . The problem now can be stated as follows: Given a set U of n elements, a collection $\{S_i\}$, $i = 1, \dots, n$ of subsets of S , and a budget L , find a subset $S' \subseteq S$ such that the total cost of elements in S' does not exceed L and the total weight of elements covered by S' is maximized. This NP-hard problem, presented by Kuller [3] in 1999, is known as the *Budgeted Maximum Coverage Problem*. The BMC problem has been used by authors such as Leskovec [72] and El-Arini [73] in their work on monitoring the blogosphere, and Curtis [74] in merging infections.

Algorithm 1, which is the greedy algorithm proposed in [3] to solve the BMC problem, solves this problem in three steps. Rather than buying the cheapest k samples, it acquires samples using the following criteria. The first step uses the sub-routine *GREEDY AREA PRICE*(S), which takes as input the collection S and selects at each step a set S_i that maximizes the ratio W'_i/c_i . The second step uses the sub-routine *GREEDY AREA*(S), which selects at each step a set S_j that maximizes W'_j constrained to the budget. The collections G and G' returned by the first and second sub-routines are the candidates for the final output. Finally, the last step uses the sub-routine *GBMC*(\cdot), which outputs the candidate solution having the maximum weight.

Clearly, the weight or number elements covered by the collections G and G' depends on R . When R tends to zero, the number of elements of S_i tends to 1 and the expression W'_i/c_i becomes $1/c_i$, which reaches its maximum when c_i reaches its minimum. In this case, the RADP-VPC-RC and GIA systems are equivalents. However, as the length of the radius is increased, more samples are included in $disk_i$, and once a sample provided by u_i is purchased all the samples within a distance less than or equal to R are considered redundant and are discarded as candidates to be acquired. Section 2.4 presents an analysis about the influence of radius R over the number of active participants.

3.4 Budget

One of the main assumptions made in most previous works is that of an infinite budget to tune the auctions. This assumption is not realistic in practice, as data collectors may have budget constraints. In the proposed mechanism, the system runs with a limited budget. Here a global budget is evenly

Algorithm 1: The Greedy Budgeted Maximum Coverage Algorithm for GIA.

input : S a collection of sets made up by the user locations

output: $S' \subseteq S$, covering set

GREEDY AREA PRICE(S)

begin

$G \leftarrow \emptyset$

$C \leftarrow 0$

$U \leftarrow S$

while $U \neq \emptyset$ **do**

 select $S_i \in U$ that maximizes $\frac{W'_i}{c_i}$

if $C + c_i \leq L$ **then**

$G \leftarrow G \cup S_i$

$C \leftarrow C + c_i$

end

$U \leftarrow U \setminus S_i$

end

return G

end

GREEDY AREA(S)

begin

$G' \leftarrow \emptyset$

$C' \leftarrow 0$

$U' \leftarrow S$

while $U' \neq \emptyset$ **do**

 select $S_j \in U'$ that maximizes W'_j

if $C' + c'_j \leq L$ **then**

$G' \leftarrow G' \cup S_j$

$C' \leftarrow C' + c'_j$

end

$U' \leftarrow U' \setminus S_j$

end

return G'

end

GBMC()

begin

$G \leftarrow$ *GREEDY AREA PRICE*(S)

$G' \leftarrow$ *GREEDY AREA*(S)

if ($w(G) \geq w(G')$) **then**

$S' = G$

else

$S' = G'$

end

end

divided by the number of rounds or auctions. However, at the end of each round r_i , if the total amount of budget assigned to round r_i is not used completely, then the remaining part is added to round $r_i + 1$.

3.5 Advantages of the GBMC Algorithm

This section illustrates the advantages in terms of coverage of using the GBMC compared with other popular algorithms such as the Greedy Set Cover (GSC) algorithm Presented by Klemberg and Tardos [75], which maximizes W'_i/c_i and corresponds to the sub-routine *GREEDY AREA PRICE* of the Algorithm 1.

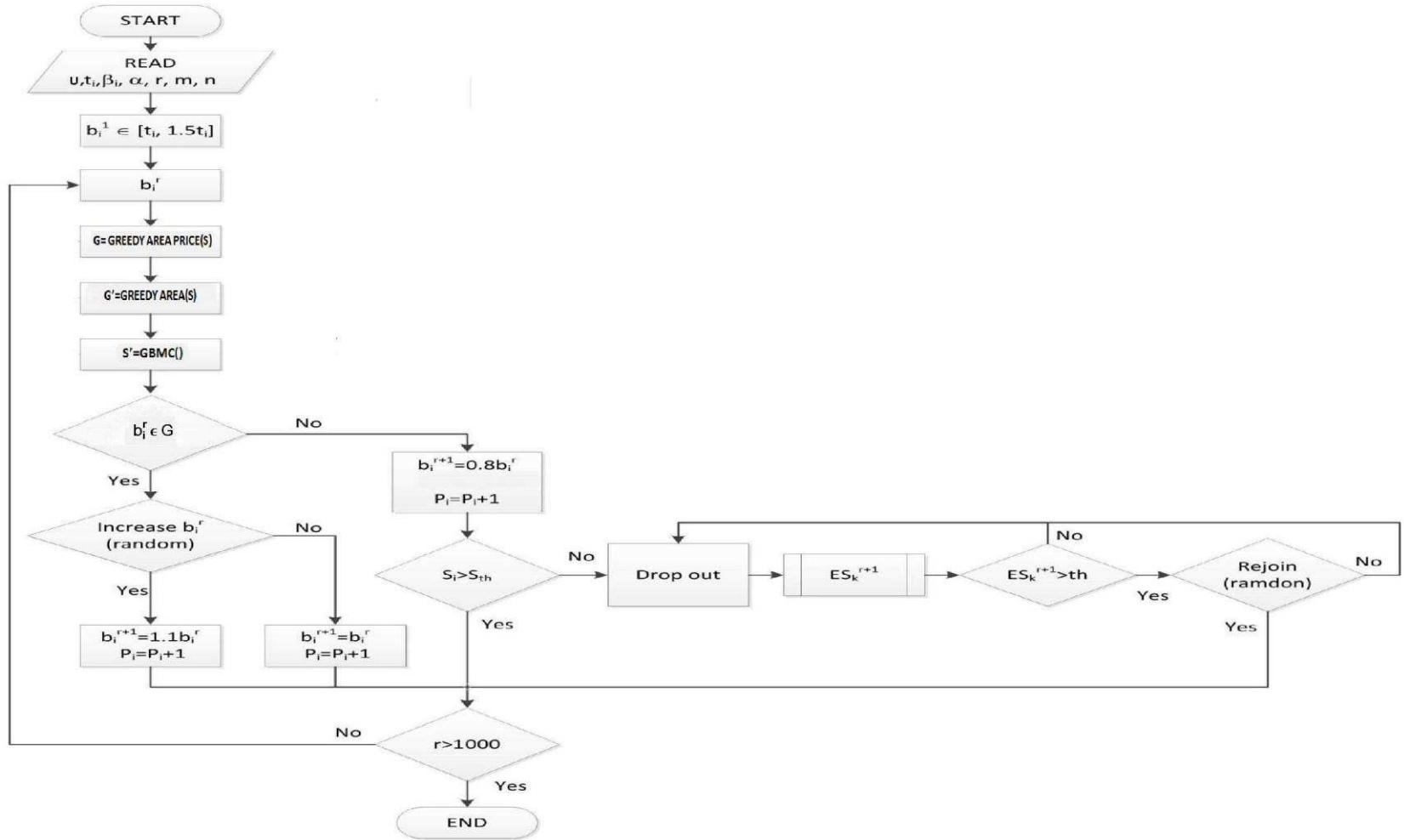


Figure 3.2: The Proposed Incentive Mechanism

The following example illustrates how the GBMC overcomes problems where the well-known GCS algorithm fails. Given a deployment of N users in a target area, and radius R , consider the following scenario. Two users (sensors), u_1 and u_2 , cover 1 and p users respectively including themselves, in other words, their samples have weights $W'_1 = 1$, and $W'_2 = p$. Let us assign the costs $c_1 = 1$ and $c_2 = p + 1$ for these samples and let us define the total budget $L = p + 1$. In this situation, Algorithm 1 acquires u_2 whereas the GSC algorithm purchases u_1 . Remember, that the optimal solution consists of covering the maximum number of participants given a budget. Let $p = 10$. In this case, GSC tries to maximize W'_i/c_i for $i = 1, 2$, which results in the purchase of u_1 because $W'_1/c_1 > W'_2/c_2$ given that $W'_1 = 1$, $c_1 = 1$, $W'_2 = 10$, $c_2 = 10 + 1$. However, in the second round there is not enough budget to acquire u_2 because one unit has been spent in the first round and the remaining 10 units of budget are not enough to acquire u_2 . In contrast, by using the budgeted approach, GBMC first uses the sub-routine *GREEDY AREA PRICE* that returns the decision to acquire u_1 , as the GSC algorithm does; however, the final decision is postponed until it compares it with the decision returned by the *GREEDY AREA* sub-routine. *GREEDY AREA*, on the other hand, acquires the sample u_i that maximizes the number of elements covered W'_i without violating the budget for $i = 1, 2$. In this case *GREEDY AREA* acquires u_2 because $W'_2 > W'_1$. The final decision is taken by the final routine GBMC() that takes as input the results returned by the two prior sub-routines and decides which of the results will be acquired comparing the number of elements covered by each one. So, in this case, the final decision is to purchase u_2 . Although in this case there is not enough budget to buy the other sample in the next round either, it is important to emphasize that the purchasing entity will cover the same area but will be left with more money. Nonetheless, when there is enough money to buy many samples in many rounds, this three step process will show its performance advantage over the other one, in particular, in terms of coverage. This advantage will be shown later in the thesis.

CHAPTER 4

IMPLEMENTATION AND EXPERIMENTAL RESULTS

4.1 Modeling the PS System Components

This chapter describes how the different parts of the PS system were modeled and presents the results of the performance evaluation comparing the proposed mechanism versus RADP-VPC-RC. Part of this work was published in [1] (Copyright © 2012 IEEE)

Table 4.1: Simulation Parameters Copyright © 2012 IEEE [1]

Parameters	Experiment 1	Experiment 2	Experiment 3
Deployment area	100m x 100m		
Instances	100		
Deployment distribution	uniform		
Deployment distribution			Four normal distributions with parameters: $\mu_1 = (30, 80)$, $\mu_2 = (80, 80)$, $\mu_3 = (50, 50)$, $\mu_4 = (90, 30)$, $Cov = \begin{pmatrix} 40 & 0 \\ 0 & 40 \end{pmatrix}$
Uniform true valuation distribution	[0,10]		no
Normal true valuation distribution	$\mu = 5, \sigma = 2$		yes
Exponential true valuation distribution	$\mu = 5$		no
Normal true valuation distribution			$\mu_1 = 5, \mu_2 = 10, \mu_3 = 15, \mu_4 = 20, \sigma = 2$
RADP-VPC-RC	no	yes	yes
GIA	yes	yes	yes
Radius R	1:10	5	5
Budget per round	100	20:200	0:350
Beta	(3,7)	(3,7)	(3,7)
Alpha	not used	7	not used
ROI Threshold	0.5	0.5	0.5

The GIA incentive mechanism, which is shown as a flowchart in Figure 3.2, works in rounds, as follows. Initially, a number of users are randomly distributed over the area of interest. For this, we used 2-dimensional uniform and normal distributions with the parameters shown in Table 4.1. Those users who are located in the area of interest receive a message announcing that a new round is about to start. Users interested in the auction set their true valuations t_i , which are randomly

generated. We perform experiments using a normal, uniform, or exponential distribution with the parameters specified in Table 4.1. For each t_i , a bid price b_i is randomly generated from a uniform distribution between t_i and $1.5t_i$. Then, the system generates a set of user locations, which are associated to the bid prices. With this information, a file indexed by user IDs is generated to allow us to track the behavior of the users through the different rounds. This file serves as the input to Algorithm 1 that determines which and how many samples will be purchased given the bid price, the location of the users, and the budget per round. At this point, Algorithm 1 chooses the winners of the round, those users to whom the system will buy the samples from. Subsequently, 50% of those winners, randomly selected, increase their bid prices by 10% and all the losers decrease their bid price by 20%. Then, all the participants increase their participation counter by 1. Next, the losers evaluate their Return of Investment (ROI) index using Equation 2.1. If this value is above a pre-defined threshold, equal to 0.5 in this work, they remain participating in the auction; otherwise, they drop out. At this point, the system sends the highest price paid per sample and the locations that were not covered in the previous round to those users who dropped out of the system in previous rounds. Using this privileged information, these users (losers) evaluate their chances of winning in the next round evaluating his ROI function for the next round using Equation 2.2. If this value is above the threshold, they rejoin the system with a 50% probability; otherwise, they remain out of the auctions. Also, users could use the location information to move to different areas and increase their chances to win the auction. This process continues and losers constantly receive this privileged information and test their chances to rejoin. The entire process repeats until a series of n different rounds are carried out, which in our case $n = 1000$, or the data collector runs out of money. It is worth mentioning that, for comparison purposes, the values utilized in all the parameters described in this paragraph are the same ones used by the authors of RADP-VPC-RC in [2].

Another part of the system that needs to be modeled in the simulation tool is the user mobility. As stated before, we do not assume static users. After a random initial deployment, the mobility algorithm computes the locations of the participants for the next round using a displacement vector (\vec{R}) that contains their displacement from the previous round. This vector is defined as the addition of a displacement vector on the X axis (\vec{X}) and another vector on the Y axis (\vec{Y}), meaning

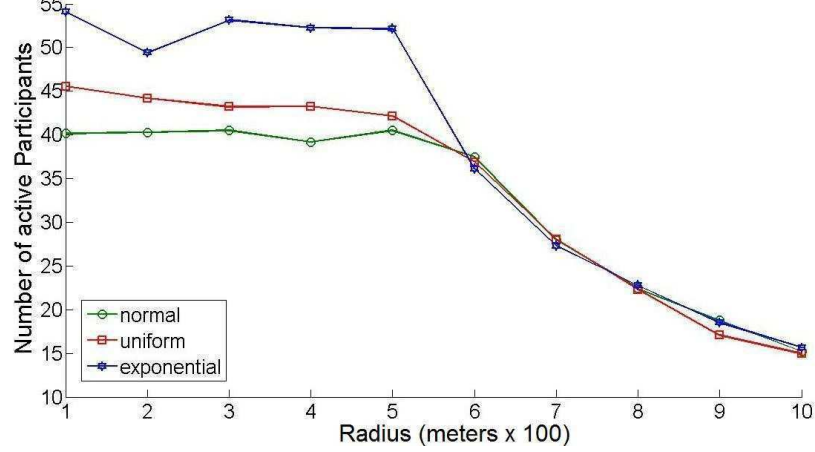


Figure 4.1: Radius vs. Number of Active Participants Copyright © 2012 IEEE [1]

$\vec{R} = \vec{X} + \vec{Y}$. The displacement vector on each axis is generated using a random function (i.e., $(-1)^{uni(0,2)}$) that defines the mobility direction. Note that the magnitude of the vector represents the participant's speed, which value is generated as the multiplication of the number of steps, randomly generated between 0 and 3, times the step length, randomly generated between 0 and 11.

4.2 Performance Evaluation

Three sets of experiments are included. The first set is meant to determine the ideal length of radius R . Once an ideal R is found, the second set consists of experiments to compare the average number of active participants, and average cost per round between the RADP-VPC-RC and GIA algorithms using different distributions for user true valuations. Finally, the third set of experiments is meant to demonstrate the superiority of the GIA algorithm in terms of coverage under conditions of stratified locations of the participants in the area of interest.

4.2.1 Experiment 1: Determining the Ideal Length of Radius R

In order to study the relationship between the radius R and the number of active participants and the average price per round, the length of R was increased from 1 to 10 in steps of one. One thousand rounds were executed and the experiment was repeated 50 times and the results averaged using a fixed budget of 100 units per round. For every length of R , the average number of active

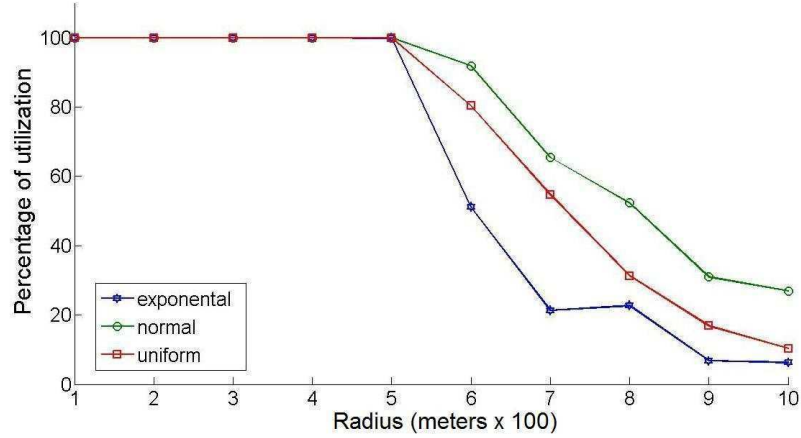


Figure 4.2: Radius vs. Budget Assigned per Round Copyright © 2012 IEEE [1]

participants was recorded. This process was carried out using three different distributions for the true valuations of the users and the uniform distribution for the initial deployment of the participants. The idea behind this experiment is to find the length of the radius R for which the average number of active participants decreases significantly. Figure 4.1 shows that when the radius R is increased beyond 5, the number of active participants decreases regardless of the distribution. This is based on the following rationale. If the radius is set as long as the distance from the center of the user to the farthest point, the algorithm will only buy one sample because this disk will cover the entire population. However, from the sensing point of view, the sensed data might not be a good representation of the real values in all places. On the other hand, if the radius is too small, the algorithm will buy more samples than necessary, as the users are too close to each other and the data will present a high level of redundancy. In short, there is a tradeoff between the radius R and the granularity and quality of the sensed data.

Figure 4.2 shows that if a radius bigger than 5 is used, the GIA algorithm does not use all the allocated budget per round to buy samples. However, when $R \leq 5$ the algorithm still uses the entire budget per round. Putting the results of these two figures together, turns out that value of $R = 5$ provides a good level of participation while rendering non redundant data. In practice, however, the value of R has to be set to a number that considers how the variable of interest changes with the distance.

4.2.2 Experiment 2: Comparing the Performance Metrics

The main goal of this experiment is to determine the number of active users in the system as the budget per round is increased from 20 to 200 when the users set their true valuations using a normal distribution with $\mu = 5$ and $\sigma = 2$, an exponential distribution with $\mu = 5$, and a uniform distribution between $[0,10]$. In the case of the GIA algorithm, a radius $R = 5$ was used.

Figure 4.3(a) shows that both algorithms present a very similar performance buying a similar number of samples per round per budget. This result basically says that both algorithms, GIA and RADP-VPC-RC, include effective methods for user participation and rejoining the system. Although the number of active users are similar, the GIA algorithm buys samples in a different manner since it avoids to buy samples from close by places that will render fairly similar or redundant data.

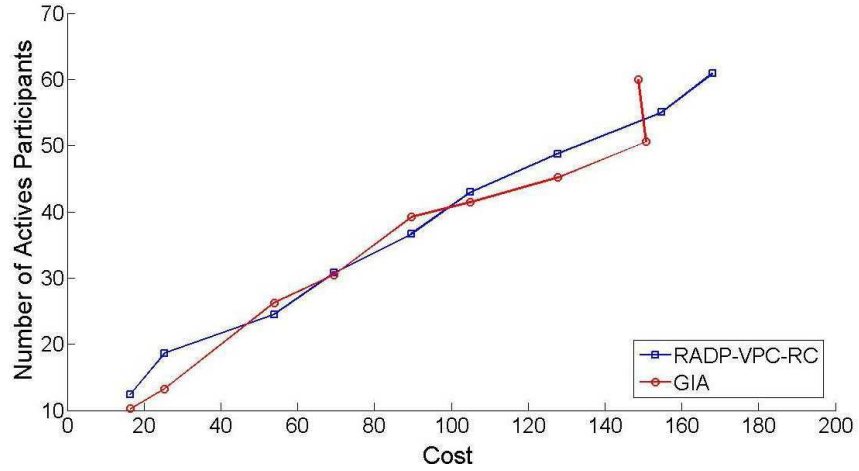
4.2.3 Experiment 3: Comparing Coverage

The main goal of this experiment is to compare the amount of the area of interest covered by the GIA and RADP-VPC-RC algorithms. For this purpose a stratified scenario was simulated. This scenario could correspond to the case where participants from different socioeconomic status having different true valuations live in distinct regions of the area (city). In this scenario the RADP-VPC-RC algorithm does not choose a representative sample of the population and instead, purchases data containing a high level of redundancy. On the other hand, GIA is able to buy non-redundant samples from each of the user types. The initial user deployment of locations was generated with a cluster structure. Four clusters with $\mu_1 = (30, 80)$, $\mu_2 = (80, 80)$, $\mu_3 = (50, 50)$, and $\mu_4 = (90, 30)$ respectively, and covariance matrix

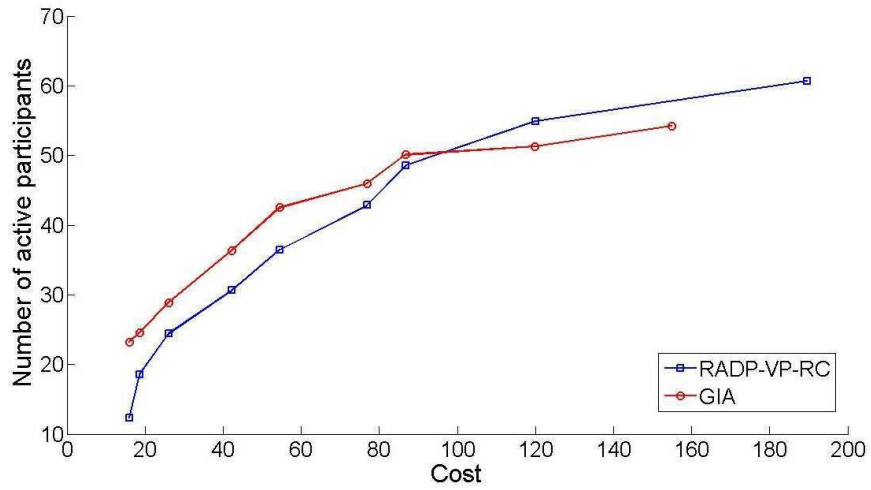
$$Cov = \begin{pmatrix} 40 & 0 \\ 0 & 40 \end{pmatrix}$$

were generated with 25 locations each, for a total of 100 locations. The true valuations of the users were generated using normal distributions with $\mu_1 = 5$, $\mu_2 = 10$, $\mu_3 = 15$, $\mu_4 = 20$, and $\sigma = 2$. Each cluster of true valuation was associated with a cluster of location in order to simulate the expected scenario.

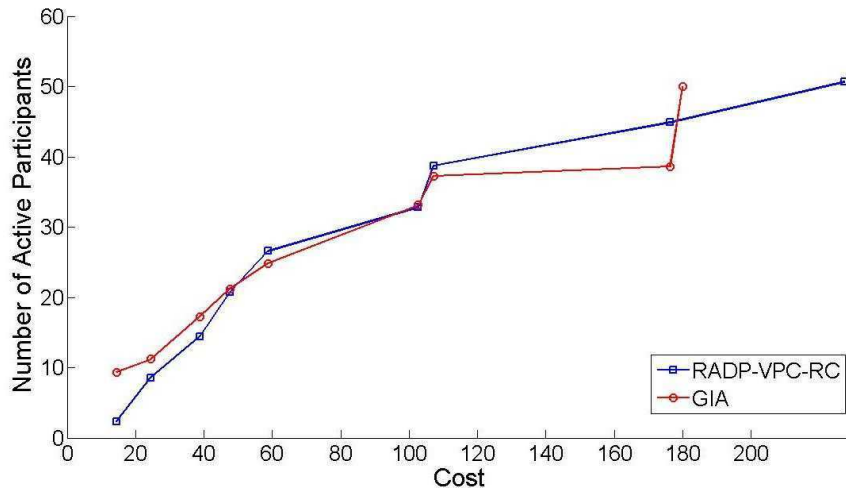
Figure 4.4 shows the percentage of area covered, the number of active participants, and total cost per round for the two algorithms as the number of samples acquired varies from 5 to 50 in steps of 5. From the Figures 4.4(a) and 4.4(b), it is clear that both algorithms have a similar number of active participants and spend a similar amount of money to buy their samples; however, as it is shown in Figure 4.4(c), it is also clear that the GIA algorithm buys those samples from users who are more spread throughout the entire area of interest, thus providing a better coverage and therefore reducing the acquisition of redundant data. For example, this last figure shows that when fifty samples are acquired, the GIA algorithm covers up to 64% more area than RADP-VPC-RC.



(a) Normal distribution.

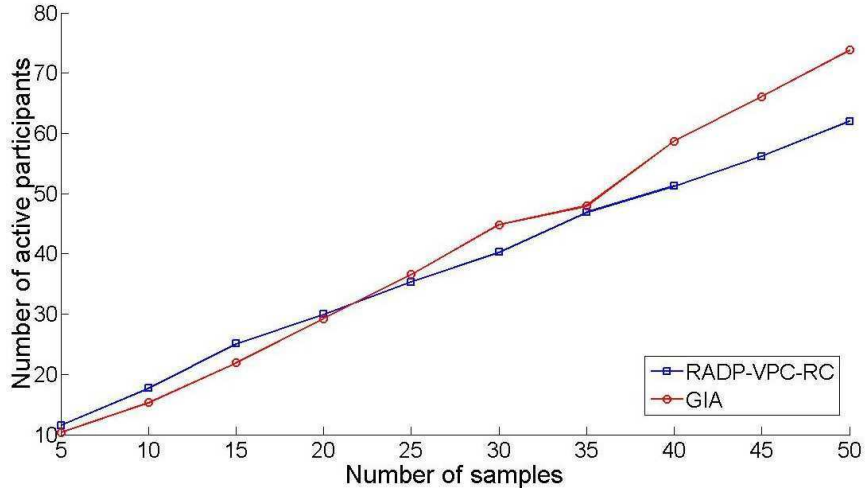


(b) Exponential distribution.

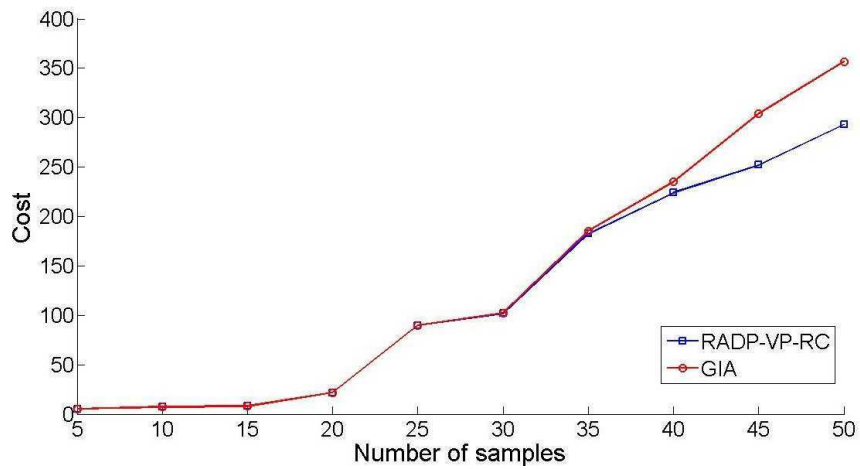


(c) Uniform distribution.

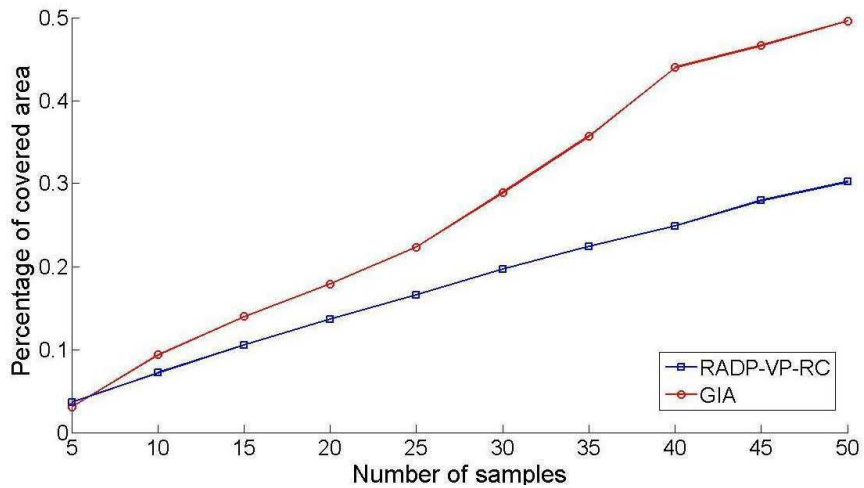
Figure 4.3: Experiment 2 - GIA vs. RADP-VPC-RC Copyright © 2012 IEEE [1]



(a) Number of Samples vs. Number of Active Participants



(b) Number of Samples vs. Cost



(c) Number of Samples vs. Percentage of Covered Area

Figure 4.4: Experiment 3 - GIA vs. RADP-VPC-RC Copyright © 2012 IEEE [1]

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusions

This thesis presents the Greedy Incentive Algorithm (GIA), a location-based reverse auction incentive mechanism for participatory sensing systems with budget constraints. The algorithm uses the location of the users to improve the coverage of the area of interest and reduce the collection of redundant data. Simulation results show the superiority of the GIA algorithm compared with RADP-VPC-RC, another reverse auction mechanism found in the literature. In general, the GIA algorithm is shown to increase the area coverage by up to 64% while maintaining a similar amount of active users in the system and spending a similar budget than the RADP-VPC-RC algorithm.

Future work is under way to model the dynamic admission of new users, as the continuous admission of new members in any round might be useful to find an equilibrium among the average number of active participants, fair prices paid per the sample, and the budget constraints. Future work include the use of Game Theory to further analyze these systems. Other future works include looking at more realistic models to simulate R and the mobility of the users as well as different ways to establish the true price and ROI for different users.

This is also the first work that survey the state of the art of mechanism of incentives for PS. Additional to GIA, this thesis include the following contributions. First, the role of PS in a emerging variety of similar concepts that share several characteristics with PS and identifies the relationship of PS with each of them. Second, a set of design issues are proposed which are used as metrics to evaluate incentive mechanisms for PS. Third, a tree-level taxonomy is proposed base on the type of user motivation used for the current PS incentive mechanisms. Fourth, a set recommendations and

guidelines are proposed for future research. Finally, the main components of a sensing data market architecture that supports the three types of PS models reviewed in this survey are proposed.

5.2 Future Research

After having reviewed the state of the art in terms of incentives mechanisms for PS, some ideas and research guidelines are proposed for future research.

5.2.1 Costs and Utility Functions

It is necessary to propose functions to estimate the cost of data acquisition base on factor such as: true valuation T , data level of perturbation for location privacy purposes P ; quality of the sensor q used for the data acquisition; elapsed *time* since the data request was done and the sample was taken; user's reputation R ; user's location (e.g., coverage) L ; and finally, taking inspiration from [33] the data's appraisal A as showed in equation 2.5 as a criteria of data acquisition.

$$C : (T, P, q, time, R, L, A) \rightarrow \Re \quad (5.1)$$

5.2.2 Schemes that Support Different Models of Sensing

In Section 2.3.1 four proposals were reviewed aimed to three different types of sensing model. The first model consider sense a region continuously in fix time intervals (i.e., rounds) some potential application are air pollution and temperature among others.. The second model considered a sensing on demand (i.e., a user has a demand of information about a specific place and the query is directed to a group of users within the target area). Finally, and less frequent scenario (e.g., natural disasters) in which is common the use TDN-like infrastructure. Here the communication between the PS system and the users (i.e., data collectors) pass through a sink, and data's trade is encouraged to maximize the users' profit. There is not any incentive mechanism that support these three types of sensing models.

5.2.3 Type of Incentive Mechanism

The literature review showed that even the potential of gaming as a powerful motivator for user participation in PS, only the work of Han *et al.* [26] have applied this mechanism to improve participation in PS. Further other types of incentive schemes such as lotteries and gambling have not been explore so far. On the other hand, although several experiments measure the level of dropping out and retention this phenomenon only has been studied rigorously by Lee and Hoh [2] and Musthag *et a.* [30]. The former in the context of recurrent reverse auctions and latter in the context of the use of Uniform, Variable and Hidden micro-payments.

5.2.4 Working Autonomous, Intelligent Agents

Despite the abundant information in this field the literature review did not find any work in which the incentive system works on a platform intelligent agents. However several of schemes reviewed may be easily extended to this environment.

5.2.5 Economic Models

The reviewed of the state of the art showed that there are only two proposals of data markets, the works of Liu *et a.* [17] and Lee and Hoh [2] respectively. The former is based on micro-economics and includes the participation of an auctioneer (i.e., PS as entity) and n bidders (i.e., users), besides includes a complete mechanism of incentives, allocation of resources, assignation of rewards, and control of dropping out and retention. The latter, include a dynamic resource allocation and rewards as well as the use of energy as mechanism to decrease the users' reservation wage. Furthermore, factors such as QoI satisfaction and network operator's profits are elements to compute the user's rewards.

5.2.5.1 Proposal for a Data Market

Taking inspiration of the types of sensing model reviewed in section 5.2.2 as well as the works of Liu *et a.* [17], Lee and Hoh [2], and Xie *et al.* [33] this thesis proposes a data market that involves the following elements: sensor, sensor's owners, data brokers, sensor's manufacturers, a

PS entity, and a PNP. The following is the description of the participants. Sensors (e.g., those embedded in smartphones, or static sensor networks deployed in the area). Sensor's owner i.e., the carrier of a smartphone or the sensor network's owners. Sensing data brokers i.e., an entity that groups a set of sensor's owners (e.g., PNP, CS project, or just an entity which has assigned a set of sensor's owners) in certain area based on some criteria (e.g., level of user privacy, sensor quality, etc.). Sensor's manufacturers this group may represents the hardware manufacturer. The following is a description of the propose model. There is a querier that require for contextual information of a site k , the querier submit a query q to the PS system (e.g., which may be part of PNP) and it in turns re-submit the query to a set of $\{1 \cdots n\}$ data brokers which may provide the data. Data brokers compute both the cost of the query's answer and the level of quality that they may offer and then they deliver this information back to the PS System. The PS system selects the data broker that offers the best trade off among price, quality and coverage. Once the data broker is selected, it in turns outsource the sensor's owners by the use of the most convenient incentive mechanism that maximizes the group utility. The data obtained from from the sensors' owners are fused by the data broker in order to offer the level of QoI required. On the other hand, sensors' owners may provide data to different data brokers and data brokers in turn may provide data to different PS systems. In this context, as Xie *et al.* [33] propose, two sensor's owners a_i, b_j who are leaving the influence region of their corresponding data broker i and j , may wish to trade their data in order to maximize their profits; within a time window while information is still valuable, or store it for other type of clients. In a second scenario, when is required to sense a wide target area continuously (e.g., temperature, humidity, pollution levels, etc) the PS system may use a reverse auction to select the subset of data brokers that can offer the best trade of among price, quality and coverage. A third scenario in which may be necessary to use a DTN-like infrastructure, the PS may use a data trade protocol as propose by Xie *et al.* [33]. Finally, since the data's price is a function of the data's quality which in turns depends of the sensor quality; a high demand of high quality sensors will cause both the sensors manufacturers, phone's manufacturer as well as PNP will facilitate the users' access to smartphones with each time more sensing capacities, compute power, and storage

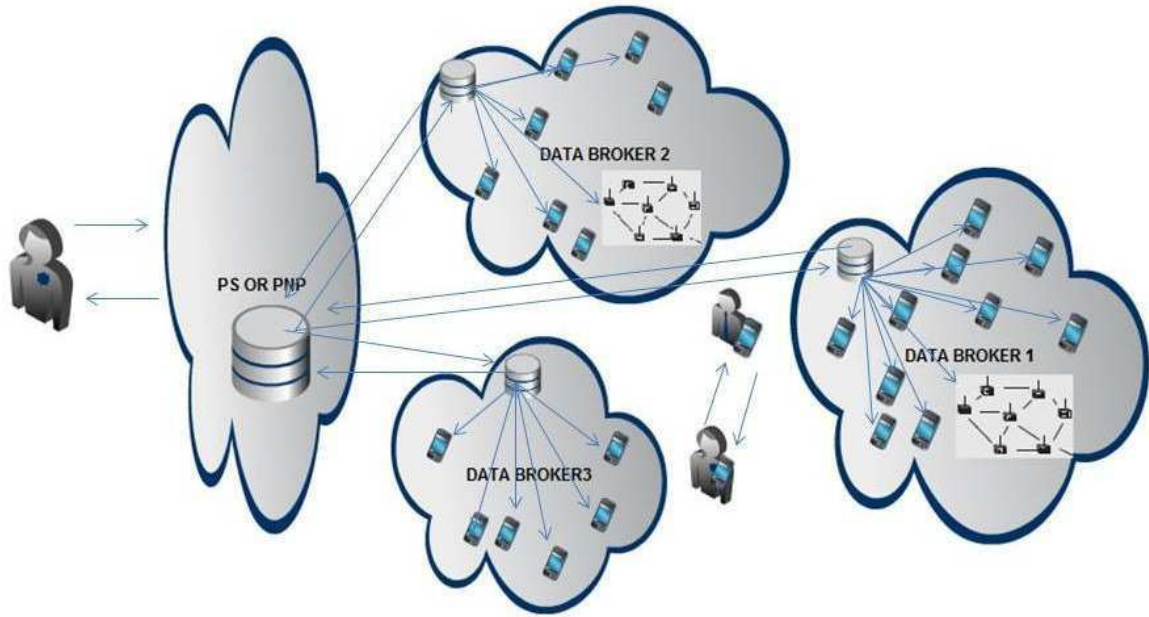


Figure 5.1: Proposal for a Incentive Architecture for PS

space for less price. Figure 5.1 sketches the main components of the three level incentive proposed architecture.

REFERENCES

- [1] L. G. Jaimes, I. J. Vergara-Laurens, and M. A. Labrador, "A location-based incentive mechanism for participatory sensing systems with budget constraints," in *PerCom*, 2012, pp. 103–108.
- [2] J.-S. Lee and B. Hoh, "Dynamic pricing incentive for participatory sensing," *Pervasive and Mobile Computing*, vol. 6, no. 6, pp. 693–708, 2010.
- [3] S. Khuller, A. Moss, and J. Naor, "The budgeted maximum coverage problem," *Inf. Process. Lett.*, vol. 70, no. 1, pp. 39–45, 1999.
- [4] E. D., "Participatory sensing: applications and architecture - internet predictions," *Internet Computing*, vol. 14, no. 1, pp. 12–42, 2010.
- [5] J. Howe, "The rise of crowdsourcing," *Wired Magazine*, vol. 14, no. 6, 2006.
- [6] "Wikipedia, <http://www.wikipedia.org>." [Online]. Available: <http://www.wikipedia.org>
- [7] M. Foth, *Handbook of Research on Urban Informatics: The Practice and Promise of the Real-Time City*. IGI Publishing, 2008.
- [8] E. Paulos, "Designing for doubt citizen science and the challenge of change," in *proceedings of Engaging Data*, 2009.
- [9] A. J. Quinn and B. B. Bederson, "Human computation: a survey and taxonomy of a growing field," in *proceedings of CHI*, 2011, pp. 1403–1412.
- [10] M.-C. Yuen, L.-J. Chen, and I. King, "A survey of human computation systems," in *proceedings of CSE*, 2009, pp. 723–728.
- [11] A. Doan, R. Ramakrishnan, and A. Y. Halevy, "Crowdsourcing systems on the world-wide web," *Commun*, vol. 54, no. 4, pp. 86–96, 2011.
- [12] J. Addison, M. Centeno, and P. Portugal, "Do reservation wages really decline? some international evidence on the determinants of reservation wages," *Journal of Labor Research*, vol. 30, no. 1, pp. 1–8, 2009.
- [13] H. Yang, J. Zhang, and P. Roe, "Using reputation management in participatory sensing for data classification," *Procedia CS*, vol. 5, pp. 190–197, 2011.
- [14] K. L. Huang, S. S. Kanhere, and W. Hu, "Are you contributing trustworthy data?: the case for a reputation system in participatory sensing," in *proceedings of MSWiM*, 2010, pp. 14–22.

- [15] H. N. Pham, B. S. Sim, and H. Y. Youn, “A novel approach for selecting the participants to collect data in participatory sensing,” in *proceedings of SAINT*, 2011, pp. 50–55.
- [16] S. Reddy, J. Burke, D. Estrin, M. H. Hansen, and M. B. Srivastava, “A framework for data quality and feedback in participatory sensing,” in *proceedings of SenSys*, 2007, pp. 417–418.
- [17] C. H. Liu, P. Hui, J. W. Branch, C. Bisdikian, and B. Yang, “Efficient network management for context-aware participatory sensing,” in *proceedings of SECON*, 2011, pp. 116–124.
- [18] S. Reddy, D. Estrin, and M. B. Srivastava, “Recruitment framework for participatory sensing data collections,” in *proceedings of Pervasive*, 2010, pp. 138–155.
- [19] H. Falaki, R. Mahajan, S. Kandula, D. Lymberopoulos, R. Govindan, and D. Estrin, “Diversity in smartphone usage,” in *proceedings of MobiSys*, 2010, pp. 179–194.
- [20] K. Shilton, N. Ramanathan, S. Reddy, V. Samanta, J. Burke, D. Estrin, M. H. Hansen, and M. B. Srivastava, “Participatory design of sensing networks: strengths and challenges,” in *proceedings of PDC*, 2008, pp. 282–285.
- [21] S. Kuznetsov and E. Paulos, “Participatory sensing in public spaces: activating urban surfaces with sensor probes,” in *proceedings of Designing Interactive Systems*, 2010, pp. 21–30.
- [22] S. Gaonkar, J. Li, R. R. Choudhury, L. P. Cox, and A. Schmidt, “Micro-blog: sharing and querying content through mobile phones and social participation,” in *proceedings of MobiSys*, 2008, pp. 174–186.
- [23] N. D. Lane, S. B. E. a Mirco Musolesi, E. Miluzzo, and A. T. Campbell, “Urban sensing systems: opportunistic or participatory?” in *proceedings of HotMobile*, 2008, pp. 11–16.
- [24] A. Kapadia, N. Triandopoulos, C. Cornelius, D. Peebles, and D. Kotz, “Anonymsense: Opportunistic and privacy-preserving context collection,” in *proceedings of Pervasive*, 2008, pp. 280–297.
- [25] T. Das, P. Mohan, V. N. Padmanabhan, R. Ramjee, and A. Sharma, “Prism: platform for remote sensing using smartphones,” in *proceedings of MobiSys*, 2010, pp. 63–76.
- [26] K. Han, E. A. Graham, D. Vassallo, and D. Estrin, “Enhancing motivation in a mobile participatory sensing project through gaming,” in *proceedings of SocialCom/PASSAT*, 2011, pp. 1443–1448.
- [27] Y. Sun and K. Nakata, “An agent-based architecture for participatory sensing platform,” in *proceedings of Universal Communication*, 2010, pp. 392–400.
- [28] D. C. Parkes, R. Cavallo, F. Constantin, and S. P. Singh, “Dynamic incentive mechanisms,” *AI Magazine*, vol. 31, no. 4, pp. 79–94, 2010.
- [29] S. Reddy, D. Estrin, M. H. Hansen, and M. B. Srivastava, “Examining micro-payments for participatory sensing data collections,” in *proceedings of UbiComp*, 2010, pp. 33–36.
- [30] M. Musthag, A. Raij, D. Ganesan, S. Kumar, and S. Shiffman, “Exploring micro-incentive strategies for participant compensation in high-burden studies,” in *proceedings of Ubicomp*, 2011, pp. 435–444.

- [31] H. Muhlenbein, “The equation for response to selection and its use for prediction,” *Evolutionary Computation*, vol. 5, no. 3, pp. 303–346, 1997.
- [32] S. Kukkonen and J. Lampinen, “Gde3: the third evolution step of generalized differential evolution,” in *proceedings of Congress on Evolutionary Computation*, 2005, pp. 443–450.
- [33] X. Xie, H. Chen, and H. Wu, “Bargain-based stimulation mechanism for selfish mobile nodes in participatory sensing network,” in *proceedings of SECON*, 2009, pp. 1–9.
- [34] A. Wiggins and K. Crowston, “From conservation to crowdsourcing: A typology of citizen science,” in *proceedings of HICSS*, 2011, pp. 1–10.
- [35] H. Ben, Karen, N. Robert, and E. Edward, “Galaxy zoo: Bar lengths in nearby disk galaxies,” *Mon Not Roy Astron Soc*, vol. 415, pp. 3627–3640, 2011.
- [36] A. Wiggins, “ebirding: technology adoption and the transformation of leisure into science,” in *proceedings of iConference*, 2011, pp. 798–799.
- [37] K. Bert, *The Social Psychology of Protest*. Blackwell, 1996.
- [38] S. Bernd, L. Michael, and S. Stefan, “Collective identification and social movement participation,” *J Pers Soc Psychol*, vol. 74, no. 1, pp. 646–658, 1998.
- [39] O. Nov, O. Arazy, and D. Anderson, “Dusting for science: motivation and participation of digital citizen science volunteers,” in *proceedings of iConference*, 2011, pp. 68–74.
- [40] C. A. L. Dantec, R. G. Farrell, J. Christensen, M. Bailey, J. B. Ellis, W. A. Kellogg, and W. K. Edwards, “Publics in practice: ubiquitous computing at a shelter for homeless mothers,” in *proceedings of CHI*, 2011, pp. 1687–1696.
- [41] C. A. L. Dantec, J. Christensen, M. Bailey, R. G. Farrell, J. B. Ellis, C. M. Danis, W. A. Kellogg, and W. K. Edwards, “A tale of two publics: democratizing design at the margins,” in *proceedings of Designing Interactive Systems*, 2010, pp. 11–20.
- [42] D. Mendez, A. J. Perez, M. A. Labrador, and J. J. Marron, “P-sense: A participatory sensing system for air pollution monitoring and control,” in *proceedings of PerCom Workshops*, 2011, pp. 344–347.
- [43] E. Kanjo, “Noisespy: A real-time mobile phone platform for urban noise monitoring and mapping,” *MONET*, vol. 15, no. 4, pp. 562–574, 2010.
- [44] M. Cherubini, A. Gutierrez, R. de Oliveira, and N. Oliver, “Social tagging revamped: supporting the users’ need of self-promotion through persuasive techniques,” in *proceedings of CHI*, 2010, pp. 985–994.
- [45] I. Krontiris and F. Freiling, “Urban sensing through social networks: The tension between participation and privacy,” in *proceedings of ITWDC*, 2010, pp. 1–10.
- [46] M. Demirbas, M. A. Bayir, C. G. Akcora, Y. S. Yilmaz, and H. Ferhatosmanoglu, “Crowd-sourced sensing and collaboration using twitter,” in *proceedings of WOWMOM*, 2010, pp. 1–9.

- [47] C. Kirkhope, R. Williams, C. Catlin-Groves, S. Rees, and C. Montesanti, “Social networking for biodiversity: The beeid project,” in *proceedings of i-Society*, 2010, pp. 625–626.
- [48] H. A. G. Stafford Richard, “Eu-social science: The role of internet social networks in the collection of bee biodiversity data,” *PLoS ONE*, vol. 5, no. 12, p. e14381, 2010.
- [49] E. Miluzzo, N. D. Lane, S. B. Eisenman, and A. T. Campbell, “Cenceme - injecting sensing presence into social networking applications,” in *proceedings of EuroSSC*, 2007, pp. 1–28.
- [50] N. Maisonneuve, M. Stevens, M. E. Niessen, and L. Steels, “Noisetube: Measuring and mapping noise pollution with mobile phones,” in *proceedings of ITEE*, 2009, pp. 215–228.
- [51] T. Erickson and W. A. Kellogg, “Social translucence: an approach to designing systems that support social processes,” *Trans Comput-Hum Interact*, vol. 7, no. 1, pp. 59–83, 2000.
- [52] M. Mun, S. Reddy, K. Shilton, N. Yau, J. Burke, D. Estrin, M. H. Hansen, E. Howard, R. West, and P. Boda, “Peir, the personal environmental impact report, as a platform for participatory sensing systems research,” in *proceedings of MobiSys*, 2009, pp. 55–68.
- [53] A. Wiggins, J. Howison, and K. Crowston, “Heartbeat: Measuring active user base and potential user interest in floss projects,” in *proceedings of OSS*, 2009, pp. 94–104.
- [54] “ebird, <http://ebird.org>.” [Online]. Available: <http://ebird.org/content/ebird/>
- [55] “Igda causal games white paper.” [Online]. Available: http://archives.igda.org/casual/IGDA_Casual_Games_White_Paper_2008.pdf
- [56] L. von Ahn, R. Liu, and M. Blum, “Peekaboom: a game for locating objects in images,” in *proceedings of CHI*, 2006, pp. 55–64.
- [57] E. L. M. Law, L. von Ahn, R. B. Dannenberg, and M. Crawford, “Tagatune: A game for music and sound annotation,” in *proceedings of ISMIR*, 2007, pp. 361–364.
- [58] L. von Ahn and L. Dabbish, “Designing games with a purpose,” *Commun*, vol. 51, no. 8, pp. 58–67, 2008.
- [59] S. Hacker and L. von Ahn, “Matchin: eliciting user preferences with an online game,” in *proceedings of CHI*, 2009, pp. 1207–1216.
- [60] E. Law and L. von Ahn, “Input-agreement: a new mechanism for collecting data using human computation games,” in *proceedings of CHI*, 2009, pp. 1197–1206.
- [61] E. Law, L. von Ahn, and T. Mitchell, “Search war: a game for improving web search,” in *proceedings of KDD Workshop on Human Computation*, 2009, p. 31.
- [62] E. Law and L. von Ahn, *Human Computation*. Morgan & Claypool Publishers, 2011.
- [63] L. von Ahn and L. Dabbish, “Labeling images with a computer game,” in *proceedings of CHI*, 2004, pp. 319–326.
- [64] L. von Ahn, “Games with a purpose,” *Computer*, vol. 39, pp. 92–94, 2006.

- [65] “Google image labeler, <http://images.google.com/imagelabeler/help.html>.” [Online]. Available: <http://images.google.com/imagelabeler/help.html>
- [66] “Budburst, <http://neoninc.org/budburst>.” [Online]. Available: <http://neoninc.org/budburst/>
- [67] “Floracaching, <http://networkednaturalist.org>.” [Online]. Available: <http://networkednaturalist.org/floracaching/>
- [68] N. Bulusu, C. C. T, and K. Salil, “Participatory sensing in commerce: Using mobile camera phones to track market price dispersion,” in *proceedings of UrbanSense*, 2008.
- [69] L. Deng and L. P. Cox, “Livecompare: grocery bargain hunting through participatory sensing,” in *proceedings of HotMobile*, 2009.
- [70] Y. Dong, S. S. Kanhere, C. T. Chou, and N. Bulusu, “Automatic collection of fuel prices from a network of mobile cameras,” in *proceedings of DCOSS*, 2008, pp. 140–156.
- [71] “Mobishop: Using mobile phones for sharing consumer pricing information, in demo session of the intl. conference on distributed computing in sensor systems, santorini, greece, june 2008.” [Online]. Available: <http://www.cse.unsw.edu.au/~salilk/publication.html>
- [72] J. Leskovec, A. Krause, C. Guestrin, C. Faloutsos, J. M. VanBriesen, and N. S. Glance, “Cost-effective outbreak detection in networks,” in *proceedings of KDD*, 2007, pp. 420–429.
- [73] K. El-Arini, G. Veda, D. Shahaf, and C. Guestrin, “Turning down the noise in the blogosphere,” in *proceedings of KDD*, 2009, pp. 289–298.
- [74] D. E. Curtis, S. V. Pemmaraju, and P. Polgreen, “Budgeted maximum coverage with overlapping costs: Monitoring the emerging infections network,” in *proceedings of ALENEX*, 2010, pp. 112–123.
- [75] J. M. Kleinberg and É. Tardos, *Algorithm design*. Addison-Wesley, 2006.

APPENDICES

Appendix A: Permissions



Copyright Clearance Center **RightsLink®** [Home](#) [Create Account](#) [Help](#)

IEEE
Requesting permission to reuse content from an IEEE publication

Title: A location-based incentive mechanism for participatory sensing systems with budget constraints

Conference Proceedings: Pervasive Computing and Communications (PerCom), 2012 IEEE International Conference on

Author: Jaimes, L. G.; Vergara-Laurens, I.; Labrador, M.A.

Publisher: IEEE

Date: 19-23 March 2012

Copyright © 2012, IEEE

User ID
Password

Enable Auto Login

LOGIN

[Forgot Password/User ID?](#)

If you're a copyright.com user, you can login to RightsLink using your copyright.com credentials.

Already a RightsLink user or want to [learn more?](#)

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

- 1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
- 2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
- 3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Figure A.1: Permission

Appendix B: Common Terms and Abbreviations Used in this Thesis

Table B.1: Common Terms and Abbreviations Used in this Thesis

Users	Private owners of the sensors
PS	Participatory Sensing
CS	Citizen Science
QoI	Quality of Information
HC	Human Computing
CDS	CrowdSourcing
SC	Social Computing
CI	Collective Intelligence
RADP	Reverse Auction Dynamic Price
VPC	Virtual Participation Credit
RC	Recruitment
RADP-VPC-RC	Reverse Auction Dynamic Price with Virtual Participation Credit and Recruitment Algorithm
GBMC	Greedy Budgeted Maximum Cover
UDMA	Univariate Model Distribution Algorithm
EGDE3	Third Evolution step Generalized Differential Evolution
GA	Genetic Algorithm
PNP	Phone Network Provider
NFS	National Science Foundation
SMS	Short Message Service
VOIP	Voice over Internet Protocol
DTN	Delay Tolerant Networks
ROI	Return Of Investment
CENS	Center for Embedded Networked Sensing
CRM	Community Resource Messenger