Spring 4-11-2023

# Network Economics-based Crowdsourcing in Online Social Networks

Natasha S. Kubiak
*University of New Mexico - Main Campus*

Natasha Kubiak

*Candidate*

Electrical and Computer Engineering

*Department*

This thesis is approved, and it is acceptable in quality and form for publication:

*Approved by the Thesis Committee:*

Dr. Eirini Eleni Tsiropoulou, Chair

Dr. Jim Plusquellic, Member

Dr. Symeon Papavassiliou, Member

# Network Economics-based Crowdsourcing in Online Social Networks

by

**Natasha Kubiak**

B.S., University of New Mexico, NM, 2022

THESIS

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Master of Science
Computer Engineering

The University of New Mexico

Albuquerque, New Mexico

May, 2023

# Dedication

*To my younger sister Valeska, for giving me unending love and sisterhood.*

*Tiger got to hunt, bird got to fly; Man got to sit and wonder 'why, why, why?'*
*Tiger got to sleep, bird got to land; Man got to tell himself he understand.*

*Kurt Vonnegut, Cat's Cradle*

# Acknowledgments

I would like to thank my advisor Dr. Eirini Eleni Tsiropoulou who has guided me tremendously through my academic career. Her advice and assistance has aided me in becoming a more confident engineer, and I want to say thank you for having faith in me. Secondly, I would like to say thank you to Aisha B Rahman and Siraj Sadman for their guidance in mentoring me. Finally, I would like to express my gratitude to the committee members, Dr. Symeon Papavassiliou, Dr. Jim Plusquellic, and my advisor, Dr. Eirini Eleni Tsiropoulou, for generously sharing their valuable time and expertise.

# Network Economics-based Crowdsourcing in Online Social Networks

by

**Natasha Kubiak**

B.S., University of New Mexico, NM, 2022

M.S., Computer Engineering, University of New Mexico, 2023

## Abstract

This thesis addresses the challenge of user recruitment by various competing marketing agencies (MAs) in Online Social Networks. A labor economics approach, following the principles of contract theory, is devised to enable MAs to reveal the potential of each participating user to contribute a personalized level of quality and quantity of information to the crowdsourcing process. The MAs objective is to maximize their personal benefit, i.e., total utility obtained, given its budget. The latter optimization problem is formulated as a Generalized Colonel Blotto (GCB) game among the MAs, where each MA aims at incentivizing each user to report its information. A Pure Nash Equilibrium (PNE) is determined resulting in the optimal rewards each MA should provide to each user. The performance evaluation of the proposed approach is achieved via modeling and simulation, and numerical results are presented to reveal the benefits of the proposed crowdsourcing model under different scenarios.

**This work has been published in:**

**A. B. Rahman, M. S. Siraj, N. Kubiak, E. E. Tsiropoulou and S. Papavassiliou, "Network Economics-based Crowdsourcing in Online Social Networks," GLOBECOM 2022 - 2022 IEEE Global Communications Conference, Rio de Janeiro, Brazil, 2022, pp. 4655-4660, doi: 10.1109/GLOBECOM48099.2022.10001611 [1].**

# Contents

# List of Figures

# List of Tables

# Glossary

| | |
|---|---|
| $N$ | Set of Online Social Networks where $N = \{1, \ldots, n, \ldots, \lvert N \rvert\}$ |
| $T_n$ | Time spent on OSN's where $T \in [0, 1]$ |
| $SP_n$ | Self Presentation represented by the amount of personal posts posted where $SP_n \in [0, 1]$ |
| $AP_n$ | Action and participation representing the number of times that a user shares posts, joins events, and goes live where $AP_n \in [0, 1]$ |
| $GI_n$ | Gratifying interactions, represents showing the number of likes and in general, reactions that a user performs in OSN Where $GI_n \in [0, 1]$ |
| $AE_n$ | Advertisement Engagement where $AE_n \in [0, 1]$ |
| $f_n$ | Engagement factor where |
| $\tau_n$ | User's type |
| $q_n$ | User's quality and quantity of information where $q_n \in [0, 1]$ |
| $P_n$ | Probability derived from past user interactions |
| $k$ | Elasticity factor |
| $G$ | Colonel Blotto Game |

*Glossary*

| | |
|---|---|
| $MA$ | Set of two marketing agency's where $MA = A, B$ |
| $R^j$ | MA's budget where $R^j, j \in MA$ |
| $r_n^j$ | Reward received by the user $n$ from the MA $j$ |
| $U^j$ | The utility of each MA in the crowdsourcing process |
| $D$ | Difference between MA rewards where $R^j, j \in MA$ |

# Chapter 1

# Introduction

Crowdsourcing and online social networks (OSNs) have attracted the interest of marketing agencies (MAs), as a means of collecting useful data, by encouraging individuals to share information [2]. The MAs, such as advertising companies, web design firms, public relations experts, content providers, and others, exploit the information availability of OSNs' users via designing appropriate crowdsourcing mechanisms [3]. The main challenge that the existing crowdsourcing mechanisms face, is the quality and quantity of information that each recruited user can provide to the crowdsourcing process [4, 5]. In this thesis, we address this problem by initially introducing a labor economics mechanism to reveal the users' quality and quantity of information that can be provided to the MAs. Then, a novel decision-making framework is proposed to design an incentivization mechanism for the MAs to motivate the users to report their information [6].

## 1.1 Related Work & Motivation

Several crowdsourcing mechanisms have been introduced in the literature to harness data stemming from users in OSNs. In [7], the authors study the impact of the intrinsic rewards experienced by the users via participating in the crowdsourcing process, on the design of the extrinsic rewards provided by the crowdsourcers. An optimization problem of the crowdsourcers' utility is formulated and solved towards determining the optimal number of participating users and the corresponding optimal amount of allocated extrinsic rewards. The users' preferences, reputation, and activities are considered in [8] to create users' homogeneous groups, which can further contribute to the crowdsourcing process [9]. The proposed multi-community collaboration crowdsourcing model aggregates the collected information from the different users groups in order to derive recommendations for the MAs. Balancing conflicts in crowdsourcing mechanisms are explored in [10], where the goal is to maximize profits for both the crowdsourcer and the participants of the crowdsourcing which would result in a win-win situation. In this study, a game-theoretic approach is used to create a profit optimization model which can balance the interests of the participant and the crowdsourcer. Thus, there can be a prize system where each participant may be rewarded according to their type, which is based on the effort given into crowdsourcing.

The focus of research in [11] is the problem of user selection before a given deadline, where the value of the user's services is maximized under the constraint that the user's given reward does not exceed a given budget. Two mechanisms are proposed, OMZ and OMG which employs a multiple-stage sampling-accepting process to allocate tasks to users based on their marginal density. The goal is to satisfy a set of six objectives which include: computational efficiency, individual rationality, budget feasibility, truthfulness, consumer sovereignty, and constant competitiveness. In other terms the system must run in real-time, each user will have a non-negative utility, the budget will not be violated, each user reports their cost, there is an arrival and departure time, each user should have a winning chance, and finally it will operate close to the optimized solution.

Crowdsourcing as a business model is explored in [12], where the goal is to utilize the model to decrease costs, create value, and stimulate public interest. The proposed framework consists of four pillars, value proposition, value creation, value transfer, and value network. The proposed crowdsourcing business model would stimulate an advantage compared to traditional business models, by using community effort to share their knowledge. The proposal in [13] examines the issue of spatial crowdsourcing, in which workers are matched with nearby tasks and rewarded accordingly for completing sensing tasks. To address the challenge of minimizing travel costs while maximizing utility, the minimum-cost maximum utility assignment (MC-MUA) is proposed. The server assignment task (SAT) model is presented for scheduling unpredictable dynamic tasks, to better estimate assigned utility and distance traveled. Another crowdsourcing problem is addressed in [14], where workers often have difficulty selecting the correct task based on their skill level, and doing so will lower the success rate of completing the crowdsourcing task. Tasks are typically recommended by the winning probability and participation history [15], however, the authors propose an additional case of the arrival of new or inexperienced workers. The proposal of two different task-matching models would be able to suggest a task difficulty classification to the workers and thus improve computing time.

Social network crowdsourcing is leveraged in [16] in order to distribute crowdsourced tasks that can exhibit real-time opinions, suggestions, and emotions [17]. As more social network users come to rely on the opinions of those in their social circle, capturing real human insight is valued in making decisions online [18]. The approach includes formulating the task, sending task invitations, executing and receiving responses from users, and analyzing the crowd results. Overall, by requesting human input gathering human opinion becomes feasible through the use of online network crowdsourcing.

An incentivization mechanism is introduced in [19] following a two-stage Stackelberg game. Specifically, the MAs (leaders) announce a monetary reward, which optimizes their benefit from the crowdsourcing process to the users (followers). Based on this information, the users determine their optimal amount of invested effort, i.e., to collect and report information, to optimize their profit. A truth-inducing mechanism is introduced in [20] to incentivize the users to behave honestly in terms of reporting their information towards optimizing the social welfare of the crowdsourcing system. A three-stage Stackelberg approach is used in [21] to observe the incentive problem in Spatial Crowdsourcing (SC). The proposed incentive framework TACT for Spatial Crowdsourcing (SC), involves recruiting workers to provide traffic data for travel recommendations. The mechanism balances exploration and exploitation to determine the quality of workers and their social relationships; a three-party game is used to determine utility contribution, and TACT uses Combinatorial Multi-Armed Bandit (CMAB) and a Three-stage Hierarchical Stackelberg (THS) game to ensure efficient recruitment and maximize the utilities of all participating users. The problem of insufficient participation in budget-constrained online crowdsourcing systems is observed in [11]. A two-tiered social crowdsourcing model is proposed such that online users will recruit within their social circle and offload tasks onto their neighbors. The proposed online incentive mechanism, MTSC, consists of two steps: Agent Selection and Online Reverse Auction, which will optimize the online duration coverage and the unit influence, while the Online Reverse Auction step will select neighbors and a suitable reward.

A Stackelberg Game-Theoretic model is used in [22] to simulate interactions between online retailers and suppliers. The focus is on optimizing assortment planning for online retailers, using a comprehensive model to study interactions in the supply chain. The model takes into account the power imbalance of small sized retailers, where the supplier is a leader that determines a fixed cost for a variety in the assortment, and a variable cost for unit production; the retailer will react by optimizing their assortment size based on the supplier. Resource management using a distributed game-theoretic approach is observed in [23] to address scalability in centralized resource management systems. Two methods are approached, a cache competition game and main processor and co-processor congestion game. The auction-based model presented is resistant to strategic manipulation, as no application can benefit from overbidding or underbidding the true value of the resource. A dynamic Colonel Blotto Game (CBG) [24] is studied in [25], where one player is the learner with a limited budget to allocate; the learner strategically distributes the budget among different battlefields based on past observations and is playing against an adversary whose strategy is unknown. The player's goal is to reduce regret, defined as the discrepancy between the payoff obtained from the best-mixed strategy and the actual payoff achieved through the implementation of a learning algorithm [26]. A combined algorithm called LagrangeBwK-Edge, which integrates a combinatorial bandit algorithm Edge and the bandit with knapsacks (BwK) algorithm, is introduced to address the budget-constrained dynamic Colonel Blotto Game, with a sublinear regret bound and a polynomial running time. A two-player zero-sum Markov game is proposed in [27] to determine data sharing in online social networks where users are wary of their private information. In the two-player zero-sum Markov game, OSN users wish to share with their online friends but want to keep other information private from the opponent. The opponent aims to block users from sharing data they want to share or stealing data they wish to keep hidden; where any gain to the user is a loss to the opponent and vice versa. Similarly, in [28] a game-theoretic approach is taken to analyze a user's willingness to relinquish their information to OSNs. In the model, it is assumed that social media users may choose to conceal, disclose, or fabricate false information based on the influence of their social network, which is reflected in the payoff function. Players gain a utility by sharing information, the amount shared is influenced by the quantity shared by their social circle, all while a cost is associated with divulging information.

OSN user outreach is studied in [29], where a game-theoretic approach is proposed to analyze two players participating in social media outreach. The study is as follows, two Instagram accounts compete by aiming to maximize their followers while under the constraint of maintaining a positive community with the following variables taken into account: followers, number of likes on the second-to-most-recent post, the total number of posts, total number of comments on the second-to-most-recent post, and percentage of negative comments. The cost is the percentage of negative comments, while the payoff function is the amount of outreach the account can obtain. A game-theoretic approach is used to detect communities in online social networks in [30], where each agent's goal is to maximize its utility. This study uses a greedy approach to reduce the computational complexity of the game and manage NP-hardness. Instead of calculating a Nash equilibrium, a local equilibrium is found, which ensures that an agent cannot improve its utility by making small changes to its strategy.

Lexical Link Analysis (LLA) as an unsupervised learning machine paradigm is used in [30] to uncover various layers of semantic networks that play a role in generating innovative ideas from big data. A set of LLA metrics are used to simplify finding high-value data from different data sources. Through LLA, three categories can be formed: Authoritative Themes which are the main topic in data, Emerging Themes or data has the potential to gain popularity, and finally, Anomalous Themes which diverge from the Authoritative theme but could be considered for further investigation. Adapting LLA to a game theoretic model, one player is the information provider, while the other players respond with their interest in the information provided. A novel graphical game-theoretic approach is observed in [31] to address OSN

Super Users; a category of social media users with a higher scope of influence on other OSN users. To analyze how users change their strategies, a Death-Birth (DB) strategy is proposed, where some users will choose to copy the strategy of another user, while others will remain unchanged. The DB strategy is utilized to dictate the probability of a user changing strategy, which is directly linked to their neighbor's type. A game-theoretic model is proposed in [32] to investigate and comprehend the dynamics of popularity in order to understand how to attract and retain online users' attention more effectively. The model requires the consideration of multiple factors, including the attributes of an item, the decay of its attractiveness over time, the heterogeneity of individual interests, and the influence of others' decisions [33]. Items such as memes, videos, or pictures can be interacted with by players; players have different types which correlate to the relevance to that user's preferences [34,35] and the utility is the instantaneous or future effect that interaction has.

The special category of mobile crowdsourcing requests from the participating users to physically travel to specific locations in order to collect and report information. In [36], an optimization framework is introduced to determine the optimal number of recruited users to maximize the execution quality of a crowdsourcing task given a specific incentive budget constraint for the MAs. In [37], a similar crowdsourcing mechanism is observed in curating specific wall content for users on social media platforms. As users and their friends engage in content, these interactions can be used to observe what content receives more attention. In this approach, visually similar content will be presented to the content that received a positive interaction. A similar approach is introduced in [38], where the participating users can further propagate invitations to their social friends to contribute to the crowdsourcing process [39]. The same philosophy is also followed in [40], where an epidemic model is introduced to capture the propagation of the invitations to social friends.

Despite the research efforts in the field of crowdsourcing within OSNs, the problem of identifying users who can provide high quality and quantity of information by being appropriately incentivized by the marketing agencies remains an open research challenge. A different approach observed in [41], shows how to achieve profit maximization by leveraging viral marketing in online social networks. Viral Marketing uses the power of referral from satisfied users to canvas and influences other surrounding users to participate. In this observation, each user is given a sphere of influence that a marketing agency may want to invest in, given the user's available budget and social media connections. Crowdsourcing by WoM (Word of Mouth) is explored in [42], where users recruit other users to execute crowdsourcing tasks in online social networks. As the recruited users engage in crowdsourcing tasks, other users will be solicited by WoM, creating a human-driven crowdsourcing mechanism. Similarly, we observe that each user participating in WoM crowdsourcing is aiming to maximize their utility by engaging and giving an effort. Another approach in [43] discusses a cost-effective and budget-balanced task allocation problem for WoM crowdsourcing in online social networks. In WoM crowdsourcing, the focus is on social groups participating in crowdsourcing tasks, rather than only targeting individuals. The paper proposes two heuristic algorithms, CB-greedy and CB-local, to minimize overall budget consumption and create cost-effective task allocation. The proposed approach in [44] focuses on using crowdsourcing techniques for big data veracity, which refers to the correctness of the data. The solution involves sentiment analysis, where users are tasked with tagging tweets according to the emotion they evoke. The tagged emotion is compared against a verified emotion in the dataset. This comparison is used to create a ROC curve and also to evaluate the accuracy of a Bayesian predictor trained with a trinomial function, using the verified data as a benchmark. A large-scale crowdsourcing approach for smart cities is presented in [45], by using mobile users' behavior for market analysis. The proposed system would track user trails via their mobile devices and would project that data onto a layout map of a real geo-location. The statistical data would help marketers adjust management strategies, such as reshaping rental prices for stores, traffic-flow designs, and digital advertising boards.

In [46] the authors study the content factors that drive users to engage on social media. The authors gauged users actions while using social media when presented with different content factors. When users are positively affected by advertisements it correlates with a positive amount of engagement or clicks. By noting what creates positive engagement, Marketing Agency's are able to use that information to better profit. Another way target user characteristics and interactions on social media are explored is observed in [47], where efficiently mining user data, MAs can use the customer information to create positive reactions in users which in turn will create more engagement with the Marketing Agency. Capitalizing on user data, MAs can commercialize a user's likes and avoid any dislikes, which in turn will create more effort from the user to engage with an MA. In [48], social media activity and messaging are classified into different advertising categories for social media marketing. The approach aimed to understand the relationship between different advertising strategies and how a potential consumer would react to the marketing strategy. The three types of advertising messages proposed are informative, persuasive, and transformative; the reactions to these messages would be observed using a combination of crowdsourcing and machine classification. A quantitative research approach is observed in [49], which analyzes online user data to determine if social media can influence purchasing decisions. This study aims to investigate the effects of advertising and brand awareness and how flow experience can influence purchasing interest. A personalized advertising marketing scheme is introduced in [50], where tweets are captured to help classify a social media user's preferences, activities, and hobbies to produce personalized advertising content. A personalized profile will be compiled for each user based on what products or brands are mentioned in the user's tweets. Using an AdSeeker advertisement engine in conjunction with semantic analysis, ontology mapping, and classification; the system was able to provide relevant advertising to the social network user. Techniques to effectively mine Facebook data and compile it into accurate user preference profiles are explored through different models in [51]. The different models which utilized Naïve Bayes, Artificial Neural Networks, and Support Vector Machine algorithms, transformed user behaviors into specific user preferences. Results from experiments on Facebook data indicate that these features accurately reflect user preferences, with the SVM approach being the most accurate. Precision Marketing and its effect on consumer purchases on the social media app TikTok is investigated in [52]. The proposed framework combines the 'Stimulus-Organism-Response' (S-O-R) model, which involves consumers emotions, and a consumer online shopping response model, which states that the effect of external stimuli on consumers varies based on their sensitivity to the environment they are in. Overall the study found that precision marketing on an online shopping platform can increase purchase behavior by improving the perceived value of online products and encouraging users to purchase again. Finally, the study suggests that to improve precision marketing there should be efforts into incorporating personalized services, quality content, and social network marketing measures that would increase the stimulation of online purchasing. The effects of social media influence on consumer decision-making are observed in [53]. Behavioral Advertising gives marketers and businesses a strategic advantage by leveraging users' activities such as social media usage, sites visited, and online shopping behavior to create personalized advertising. In this study, a questionnaire is given to a set of users to collect their pre and post-purchase behavior in an effort to understand how social media marketing can influence users. In [54], marketing strategies are analyzed for new media e-commerce platforms with regard to a user's personal preferences. The explored marketing algorithm that is based on users' preference has two categories: Users actively search for a product they want, or users subconsciously desire a product but do not know what they explicitly want yet. With the utilization of machine learning and data mining methods, marketing agencies can better understand consumers' desires and suggest content that could entice a reaction with a higher probability. The correlation between online social network engagement and sales is discussed in [55], where actions such as likes, comments, and shares can positively impact consumer/brand relationships. The paper makes contributions in three main spaces: Firstly, it tests the relationship between social media-based brand engagement and purchase intention, indicating whether investments in social media marketing communications have a return on investment. Secondly, it proposes Engagement Intention, to measure how users engage with brands on social media and how it builds loyalty. Lastly, it evaluates different Facebook Posts and how they impact consumer opinions towards the brand. Digital Media Marketing in conjunction with the analysis of social networking trends is used in [56] to assist Marketing Agencies in finding compatible customers. The proposed system would integrate data gathered from Facebook, Twitter, and Instagram to classify users into different categories based on their preferences. A supervised learning algorithm called the Multi-layer Perceptron (MLP) is proposed as a logistical regression classifier model with hidden layers. As social media usage continues to trend upward, guaranteed profitability becomes a concern when making an investment on user resources [57]. Firms are reluctant to invest unless they can be sure of a return on investment (ROI). The ROI is dependant on the benefit of social media investment which includes reach, improved business generation, customer loyalty, improved communication with customers, and negating dissatisfaction. Different social media project examples are analzyed in order to showcase their profitability, while demonstrating ROI in social media framework.

## 1.2 Contributions & Outline

In this thesis, a labor economics approach is introduced to enable the marketing agencies to reveal the potential of each participating user to contribute a personalized level of quality and quantity of information to the crowdsourcing process. The marketing agencies compete with each other to win the ownership over the collected information by each participating user, based on their available budget and the rewards offered to the users. In the following, we assume two types of marketing agencies, where the winner of the competition has the privilege of owing and reusing, e.g., reselling, the collected information from the participating users. The other type of marketing agencies (i.e., the one that loses the competition) can only use the collected information for personal purposes, e.g., design an advertising campaign, as it has invested fewer rewards. The objective of each marketing agency in the aforementioned process and competition is to maximize their personal benefit, i.e., utility, from the overall crowdsourcing process. The latter optimization problem is formulated and solved as a Generalized Colonel Blotto (GCB) game among the marketing agencies. A Pure Nash Equilibrium (PNE) is determined resulting in the optimal rewards that each marketing agency should provide to each user. A detailed set of numerical and comparative results are presented in order to quantify the drawbacks and benefits of the proposed network economics-based crowdsourcing framework.

The rest of the thesis is organized as follows. The system model is presented in Section 2, while Section 3 introduces the contract-theoretic model [58,59] in order to reveal the participating users' potential to provide information, following the principles of labor economics. Section 4 determines the optimal incentivization rewards provided by the marketing agencies to the users based on a GCB game-theoretic approach. A simulation-based evaluation is demonstrated in Section 5, while Section 6 concludes the thesis.

# Chapter 2

# System Model

We consider a set of OSNs users $N = \{1, \ldots, n, \ldots, |N|\}$. Based on the existing scales for measuring user engagement within the OSNs, such as the Multi-dimensional Facebook Intensity Scale (MFIS), the Social Networking Activity Intensity Scale (SNAIS), the Social Media Use Integration Scale (SMUIS), and others [60, 61], the users' activity and engagement in the OSNs can be quantified based on the following metrics: (i) self-presentation $SP_n \in [0, 1]$ capturing the number of personal posts on OSNs; (ii) action and participation $AP_n \in [0, 1]$ representing the number of times that a user shares posts, joins events, and goes live; (iii) gratifications and interactions $GI_n \in [0, 1]$ showing the number of likes and in general, reactions that a user performs in OSNs; (iv) time $T_n \in [0, 1]$ spent on OSNs; and (v) advertisement engagement $AE_n \in [0, 1]$ measuring the number of times that a user clicks on an advertisement. The above metrics are normalized in the interval $[0, 1]$ for homogeneity in the presentation. Also, the values of these metrics can be retrieved by analyzing the users' social graphs of the corresponding OSNs [60].

By exploiting the above metrics, we can define each users' activity and engagement factor $f_n = w_1 SP_n + w_2 AP_n + w_3 GI_n + w_4 T_n + w_5 AE_n, \sum_{i=1}^{5} w_i = 1$, which represents how active a user is within the OSNs. Thus, we define the users' type $\tau_n = \frac{f_n}{\sum_{n=1}^{|N|} f_n}$, which represents each users' relative engagement within OSNs with respect to the rest of the users that the MAs have recruited in the crowdsourcing process. It is highlighted that the users' social characteristics are a personalized information, which is either not revealed to the MAs by the users considering their privacy concerns or it is extremely costly and time consuming to be determined by the MAs. Also, a user of high type $\tau_n$ is expected to have the potential to provide a higher quality and quantity of information, given that the user is strongly engaged with the OSNs. We define the user's quality and quantity of information as the normalized metric $q_n \in [0, 1]$.

Furthermore, we consider two categories of marketing agencies, where $MA = \{A, B\}$ denotes their set. Each MA has a total budget $R^j$, $j \in MA$ that can invest to incentivize the users to participate in the crowdsourcing process. Specifically, each user receives a reward $r_n^j$ and $r_n^{-j}$ from the MA $A$ and $B$, respectively, to be incentivized to provide its information. If an MA provides higher reward, i.e., $r_n^j > r_n^{-j}$, then, it "wins" the ownership over the reported information from a user $n$, e.g., it can resell or reuse the information for multiple purposes. On the other hand, the MA that provides lower reward to a user can exclusively use the collected information for its personal purposes, e.g., design its own advertising campaign. In the special case of a tie, i.e., $r_n^j = r_n^{-j}$, then both MAs win ownership over the collected information.

Based on the above analysis, the utility, i.e., benefit, of each MA in the crowdsourcing process can be captured as follows:

$$U^j(\mathbf{r}^j, \mathbf{r}^{-j}, k) = \sum_{n=1}^{|N|} \frac{q_n}{\pi} \arctan[k(r_n^j - r_n^{-j})] + \frac{q_n}{2} \tag{2.1}$$

where, $\mathbf{r}^j = [r_1^j, \ldots, r_{|N|}^j]$ and $\mathbf{r}^{-j} = [r_1^{-j}, \ldots, r_{|N|}^{-j}]$ are the provided rewards to the users by the MAs. The physical meaning of the MAs' utility captures the benefit that an MA receives by investing higher rewards compared to its competitor in order to win the ownership over the crowdsourced information. It is highlighted that the MAs' utility is of sigmoidal shape, given that both MAs benefit out of the crowdsourcing process, either via owning or personally using the crowdsourced information. Please note that on the contrary, the adoption of a step-like function - though simpler in its treatment - would fail to realistically represent and capture the behavior of an actual crowdsourcing system.

# Chapter 3

# Contract Theory and Users Contribution

A crowdsourcing system is characterized by information asymmetry, meaning that the MAs are not aware in advance regarding the quality and quantity of information $q_n \forall n \in [0, 1]$, of each participating user. Thus, a process that identifies the participating users' social characteristics can substantially benefit the overall crowdsourcing process for all the involved MAs. Towards this direction, we introduce a contract-theoretic approach that bridges the gap of the information asymmetry by designing personalized rewards for each user [62, 63].

Without loss of generality, we consider $R^B < R^A$, where as mentioned before $R^A, R^B$ represent the available total monetary rewards, i.e., budget, that the MAs can allocate to the $|N|$ participating users. MA $B$ is selected to participate in the contract-theoretic process in order to reveal the users' characteristics. Specifically, the MA $B$ provides personalized contracts $\{\hat{r}_n^{B*}, q_n^*\}_{\forall n \in N}$ to the users in order to incentivize them to participate in the crowdsourcing process and start revealing their personal type $\tau_n$ via the corresponding reported quality and quantity of information

$q_n$. The MA's B provided rewards are denoted as $\hat{r}_n^{B*}$, where $r_n^{B*} = \frac{\hat{r}_n^{B*}}{\sum_{n=1}^{|N|} \hat{r}_n^{B*}} R^B$ and $\hat{r}_n^{B*} \in [0,1]$. The MA is unaware of the users' types $\tau_n, \forall n \in N$ and probabilistically estimates them with probability $P_n$, which can be derived from historical data from past interactions with the users. The MA's utility function from attempting to reveal the users' characteristics is defined as: $U^B(\mathbf{q}) = \sum_{n=1}^{|N|} P_n(q_n - \hat{r}_n^B)$. The utility function $U^B(\mathbf{q})$ represents the MA's contract-theoretic profit from attempting to reveal the users' characteristics. On the other hand, the users' utility is similarly defined as the users' profit from participating in the crowdsourcing process: $U_n(q_n) = \tau_n e(\hat{r}_n^B) - q_n$, while considering its personal characteristics, as they are captured by its type $\tau_n$. The function $e(\hat{r}_n^B)$ is an evaluation function of the received reward, which is concave and strictly increasing, with $e(\hat{r}_n^B = 0) = 0$, capturing the users' increasing satisfaction from higher received rewards.

Our goal is to determine the optimal contracts $\{\hat{r}_n^{B*}, q_n^*\}$, thus, ultimately revealing the capability of each user to provide a level of quality and quantity of information $q_n^*$ by receiving an optimal reward $\hat{r}_n^{B*}$. Towards this direction, the following optimization problem is formulated.

$$\max_{\{\hat{r}_n^{B*}, q_n^*\}_{\forall n \in N}} U^B(\mathbf{q}) = \sum_{n=1}^{|N|} P_n(q_n - \hat{r}_n^B) \tag{3.1a}$$

$$\text{s.t.} \quad \tau_n e(\hat{r}_n^B) - q_n \geq 0, \forall n \in N \quad \text{(IR)} \tag{3.1b}$$

$$\tau_n e(\hat{r}_n^B) - q_n \geq \tau_n e(\hat{r}_{n'}^B) - q_{n'} \quad \text{(IC)} \tag{3.1c}$$

$$0 \leq \hat{r}_1^B < \cdots < \hat{r}_n^B < \cdots < \hat{r}_{|N|}^B \tag{3.1d}$$

The physical meaning of the above optimization problem is that the MA $B$ aims at maximizing its profit (Eq. 3.1a) from participating in the crowdsourcing process, while guaranteeing that each participating user will experience a non-negative profit (Eq. 3.1b), i.e., individual rationality (IR) condition. Also, the optimal contracts $\{\hat{r}_n^{B*}, q_n^*\}$ should provide the higher profit to each user as compared to any other contract designed for another user (Eq. 3.1c), i.e., incentive compatibility (IC) condition, in order for the users to be incentivized to participate in the crowdsourcing process. Furthermore, the optimal contracts should be characterized by fairness, monotonicity, and rationality (Eq. 3.1d), meaning that a user of higher type $\tau_n$, i.e., more active and engaged with the OSNs, has the potential to provide higher quality and quantity of information $q_n$, thus, it should receive a higher reward $\hat{r}_n^B$, and experience a higher profit $U_n$. It is noted that for presentation purposes and without loss of generality, we have sorted the users' ID as $\tau_1 < \cdots < \tau_n < \cdots < \tau_{|N|}$.

Towards solving the non-convex optimization problem Eq. (3.1a) – Eq. (3.1d), we analyze and reduce its constraints in the following analysis. Focusing on the IR condition in Eq. (3.1b), and by exploiting the IC condition in Eq. (3.1c), we have: $\tau_n e(\hat{r}_n^B) - q_n \geq \tau_n e(\hat{r}_{n'}^B) - q_{n'} \geq \cdots \geq \tau_n e(\hat{r}_1^B) - q_1$. Thus, we conclude that $\tau_n e(\hat{r}_n^B) - q_n \geq \tau_1 e(\hat{r}_1^B) - q_1 \geq 0$, given that $\tau_n > \tau_1, \forall n \in N$. Also, the MA $B$ will provide just the sufficient reward to each user to incentivize them to participate in the crowdsourcing process. Thus, we can write the previous outcome as $\tau_1 e(\hat{r}_1^B) - q_1 = 0$, which can replace the constraint in Eq. (3.1b). The physical meaning of this analysis is that if the user with the lowest type can be incentivized to participate in the crowdsourcing process, then, all the other users, which are more active in the OSNs, will also be incentivized. Then, we focus our analysis on reducing the IC constrains (Eq. 3.1c). Towards this direction, we initially define the following terminology for the IC constraints: (a) $n, n', n' \in \{n+1, \ldots, |N|\}$ upward IC (UIC); (b) $n, n', n' \in \{1, \ldots, n-1\}$ downward IC (DIC); (c) $n, n+1, n \in N$ local UIC (LUIC); and (d) $n, n-1, n \in N$ local DIC (LDIC).

**Theorem 1.** *The LDIC constraint can represent all the DIC constraints [64].*

*Proof:* Initially, we consider three representative user types $\tau_{n-1} < \tau_n < \tau_{n+1}$, and we write the corresponding IC constraints, as follows: $\tau_{n+1}e(\hat{r}^B_{n+1}) - q_{n+1} \geq \tau_{n+1}e(\hat{r}^B_n) - q_n$ and $\tau_n e(\hat{r}^B) - q_n \geq \tau_n e(\hat{r}^B_{n-1}) - q_{n-1}$. Given the monotonicity, rationality, and fairness conditions, we have: $\tau_n > \tau_{n-1} \Leftrightarrow \hat{r}^B_n > \hat{r}^B_{n-1} \xLeftrightarrow{e\nearrow} e(\hat{r}^B_n) > e(\hat{r}^B_{n-1}) \xLeftrightarrow{\tau_{n+1}>\tau_n} \tau_{n+1}[e(\hat{r}^B_n) - e(\hat{r}^B_{n-1})] > \tau_n[e(\hat{r}^B_n) - e(\hat{r}^B_{n-1})] \geq q_n - q_{n-1}$, where the last step can be derived from the previously written IC constraints. We recursively write the previous outcome for all the contracts, and we have: $\tau_{n+1}e(\hat{r}^B_{n+1}) - q_{n+1} \geq \tau_{n+1}e(\hat{r}^B_{n-1}) - q_{n-1} \geq \cdots \geq \tau_{n+1}e(\hat{r}^B_1) - q_1$. Thus, we have shown that if $\tau_n e(\hat{r}^B_n) - q_n \geq \tau_n e(\hat{r}^B_{n-1}) - q_{n-1}$, all the DIC constraints hold true. ■

**Theorem 2.** *The LDIC constraint can represent all the UIC constraints [65].*

*Proof:* Similarly, we consider $\tau_{n-1} < \tau_n < \tau_{n+1}$, and from writing the IC constraints, we have: $\tau_{n-1}e(\hat{r}^B_{n-1}) - q_{n-1} \geq \tau_{n-1}e(\hat{r}^B_n) - q_n$ and $\tau_n[e(\hat{r}^B_{n+1}) - e(\hat{r}^B_n)] \geq \tau_{n-1}[e(\hat{r}^B_{n+1}) - e(\hat{r}^B_n)]$, where the latter step holds true given that $\tau_n > \tau_{n-1}$. By combining these inequalities, we have: $\tau_{n-1}e(\hat{r}^B_{n-1}) - q_{n-1} \geq \tau_{n-1}e(\hat{r}^B_n) - q_n \geq \tau_{n-1}e(\hat{r}^B_{n+1}) - q_{n+1}$. By recursively applying this outcome, we have: $\tau_{n-1}e(\hat{r}^B_{n-1}) - q_{n-1} \geq \tau_{n-1}e(\hat{r}^B_{n+1}) - q_{n+1} \geq \cdots \geq \tau_{n-1}e(\hat{r}^B_{|N|}) - q_{|N|}$. Thus, we conclude that if the LDIC constraint holds true, then, all the UIC constraints hold true. ■

Based on the above analysis, the optimization problem (3.1a) – (3.1d), can be rewritten with the reduced constraints as follows,

$$\max_{\{\hat{r}^B_n, q_n\}_{\forall n \in N}} U^B(\mathbf{q}) = \sum_{n=1}^{|N|} P_n(q_n - \hat{r}^B_n) \tag{3.2a}$$

$$\text{s.t.} \;\; \tau_1 e(\hat{r}^B_1) - q_1 = 0 \tag{3.2b}$$

$$\tau_n e(\hat{r}^B_n) - q_n = \tau_n e(\hat{r}^B_{n-1}) - q_{n-1} \tag{3.2c}$$

$$0 \leq \hat{r}^B_1 < \cdots < \hat{r}^B_n < \cdots < \hat{r}^B_{|N|} \tag{3.2d}$$

The above convex optimization problem can be solved based on standard convex optimization tools and derive the optimal contracts $\{\hat{r}_n^{B*}, q_n^*\}$. Thus, the MA $B$ has already determined the optimal rewards $r_n^{B*} = \frac{\hat{r}_n^{B*}}{\sum\limits_{n=1}^{|N|} \hat{r}_n^{B*}} R^B, \forall n \in N$ that should provide to the users in order to incentivize them to report their personal optimal quality and quantity of information $q_n^*, \forall n \in N$, which also optimizes the users' profit. In the following section, our goal is to determine the optimal rewards $r_n^{A*}$ that should be provided by the other MA $A$, who has competing interests compared to the MA $B$. However, at this point, the users' potential of providing different levels of quality and quantity of information $q_n^*$ have been revealed to both MAs.

# Chapter 4

# Generalized Colonel Blotto Game – A Reward Mechanism

In this section, we introduce a non-cooperative game-theoretic approach in order to determine the optimal reward vectors $\mathbf{r^{A*}} = [r_1^{A*}, \ldots, r_{|N|}^{A*}]$ and $\mathbf{r^{B*}} = [r_1^{B*}, \ldots, r_{|N|}^{B*}]$ that will be invested by both marketing agencies, given that the users' characteristics, i.e., $q_n^*, \forall n \in N$, have been revealed in the crowdsourcing system, in order to maximize their benefits from the crowdsourcing process, as captured by their utility function (Eq. 2.1). Towards this direction, the theory of Colonel Blotto (CB) game is adopted. Based on the traditional CB model, two competitors compete among each other in a set of battlefields, and the one that invests more resources wins the battlefield. The latter one is characterized by a value, i.e., importance of the battlefield. The competitor, who wins the battlefield, enjoys the battlefield's value as a trophy, while the other competitor receives zero utility, even if it invested part of its resources [66].

The traditional CB model has been relaxed in [67] by introducing the Generalized Colonel Blotto (GCB) game. Under the GCB model, the competitor, who lost the battlefield, still enjoys a level of utility, given that it invested part of its resources in the battlefield. Thus, the received benefit from the competitors is captured by a strictly increasing sigmoidal function with respect to the difference of the competitors' invested resources, similar to the one presented in Eq. (2.1). The elasticity factor $k, k \in \mathbb{R}^+$ captures the fairness in terms of the received utility from the competitors. Specifically, a greater value of the elasticity factor $k$ favors the greater utility of the winning competitor, i.e., the one that invested a higher amount of resources to the battlefield. On the other hand, smaller values of $k$ support a fairness balance among the two competitors, while still providing higher utility to the one that wins the battlefield.

The concept of the Generalized Colonel Blotto game is adopted in this thesis in order to capture the competing interests of the two involved MAs. We define the GCB game as $G = \{MA, \{\mathcal{R}^j\}_{\forall j \in MA}, \{R^j\}_{\forall j \in MA}, N, \{q_n^*\}_{\forall n \in N}, \{U^j\}_{\forall j \in MA}\}$, where $MA = \{A, B\}$ is the set of marketing agencies, i.e., competitors, $\mathcal{R}^j = \{\mathbf{r^j} | \sum_{n=1}^{|N|} r_n^j \leq R^j, r_n^j \geq 0\}$ is the strategy space for $j \in MA$, $R^j$ is the marketing agency's $j$ available budget to provide rewards to the participating users, $N$ is the set of users (representing the battlefields, as described above), $q_n^*$ is the users' quality and quantity of information, i.e., the value of the battlefield, and $U^j$ is the utility function of the marketing agency (i.e., player/competitor) $j$, as defined in Eq. (2.1). The MAs compete with each other in order to determine the optimal rewards that each one should provide to the participating users in order to maximize his/her utility $U^j$. Given the MAs competing interests, the corresponding minimax problem is formulated as follows:

$$\min_{\mathbf{r^{-j}} \in \mathcal{R}^\mathbf{j}} \max_{\mathbf{r^j} \in \mathcal{R}^\mathbf{j}} U^j(\mathbf{r^j}, \mathbf{r^{-j}}, k) = \sum_{n=1}^{|N|} \frac{q_n}{\pi} \arctan[k(r_n^j - r_n^{-j})] + \frac{q_n}{2} \tag{4.1}$$

while considering the budget constraints $R^j$ of the MAs.

Our goal is to solve the GCB game and determine a Pure Nash Equilibrium (PNE) in order to derive the optimal allocation of the MAs' rewards to the participating users.

**Theorem 3.** *The GCB game*

$$G = \{MA, \{\mathcal{R}^j\}_{\forall j \in MA}, \{R^j\}_{\forall j \in MA}, N, \{q_n^*\}_{\forall n \in N}, \{U^j\}_{\forall j \in MA}\}$$

*has a PNE* $\mathbf{r}^{\mathbf{A}*}, \mathbf{r}^{\mathbf{B}*}$, *where* $\mathbf{r}^{\mathbf{B}*} = [r_1^{B*}, \dots, r_{|N|}^{B*}]$ *and* $\mathbf{r}^{\mathbf{A}*} = \mathbf{r}^{\mathbf{B}*} + [z_1^*, \dots, z_{|N|}^*]$, *where* $r_n^{B*}$ *is derived by the solution of the optimization problem (3.2a) - (3.2d),* $z_{|N|}^*$ *is the positive solution of* $z_{|N|} + \sum_{n=1}^{|N|-1} \sqrt{\frac{1}{k^2 q_{|N|}}(k^2 z_{|N|}^2 q_n + q_n - q_{|N|})} = D$, *and* $z_n^* = \sqrt{z_{|N|}^{*2} \frac{q_n}{q_{|N|}} + \frac{q_n - q_{|N|}}{q_{|N|} k^2}}, \forall n \in N \backslash \{|N|\}$ *for* $k \geq \max\{\frac{1}{R^A - R^B} \frac{|N|-1}{\sqrt{q_{|N|}(2|N|-1)}}, \frac{1}{R^A - R^B} \sum_{n=1}^{|N|-1} \sqrt{\frac{q_n - q_{|N|}}{q_{|N|}}}\}$, *with* $R^A > R^B$.

*Proof:* The proof of Theorem 3 can be derived by following similar analysis as in [67]. Due to space limitations, the proof is omitted here. ■

Given the above analysis, the optimal rewards $\mathbf{r}^{\mathbf{A}*}$ and $\mathbf{r}^{\mathbf{B}*}$ that the competing marketing agencies should provide to the participating users are determined, while both marketing agencies maximize their utilities given their budgets.

# Chapter 5

# Numerical Results

In this section, a detailed numerical, market-based, evaluation is presented to demonstrate the characteristics of the proposed network economics-based crowdsourcing framework. We consider two competing marketing agencies with $R^A > R^B$. The Facebook graphs of $|N| = 10$ indicative users have been derived for an one-month time period and their corresponding metrics have been normalized, i.e., $SP_n, AP_n, GI_n, T_n, AE_n \in [0,1], \forall n \in N$. The weights of the users' social characteristics are $\mathbf{w} = [0.05, 0.3, 0.15, 0.05, 0.45]$. The elasticity factor is $k = 10$, $R^A = 308, R^B = 300$ monetary units, and the probability of the users' types is derived from a uniform distribution, unless otherwise explicitly stated.

## 5.1 Performance Evaluation & Comparative Analysis

In this section, we present some comparative results to show the pure operation and performance of the proposed framework, as well as its benefits. We consider the

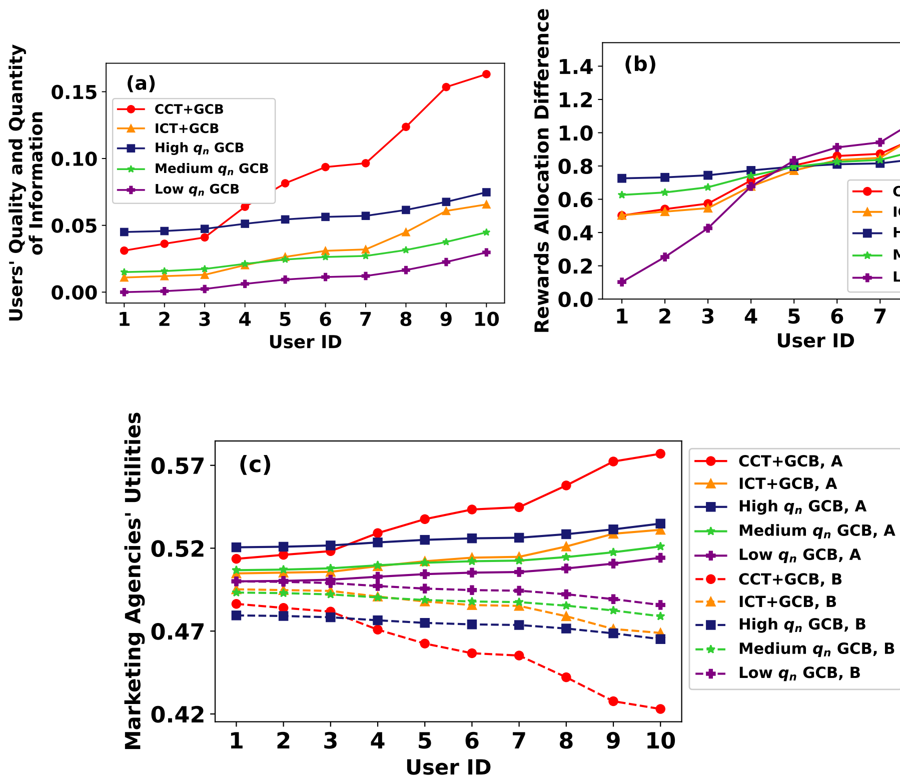


Figure 5.1: Performance evaluation and comparative analysis.

following five comparative scenarios: (i) incomplete contract-theoretic GCB model (ICT+GCB), where information asymmetry exists in the crowdsourcing system, and the users' quality and quantity of information values $q_n^*, \forall n \in N$ are derived as proposed in Section 3, while the MAs' rewards are determined based on the GCB model (Section 4); (ii) complete contract-theoretic GCB model (CCT+GCB), where the MAs are assumed to have complete information regarding the users' types; and (iii)–(v) High, Medium, and Low $q_n^*$ GCB, where the MAs are assumed to have a predetermined high, medium, and low estimation of the users' quality and quantity of information $q_n^*$, respectively.

Fig. 5.1a – Fig. 5.1c present the users' quality and quantity of information $q_n^*$, the MAs' rewards allocation difference $r_n^{A*} - r_n^{B*}$, and their achieved utility $U^A$ and $U^B = 1 - U^A$, as a function of the users' ID, respectively. The results show that under the complete information scenario (CCT+GCB), the marketing agencies can fully exploit the participating users' social characteristics and collect high levels of quality and quantity of information as compared to the information asymmetry (ICT+GCB) scenario, and the static scenarios where the marketing agencies probabilistically or statically derive the users' characteristics, respectively (Fig. 5.1a). Also, we observe that the more valuable is a user in terms of providing crowdsourced information (i.e., the higher its ID is in our evaluation), the higher is the corresponding reward allocation difference (Fig. 5.1b), as the MA with large budget competes more aggressively to collect information from the more valuable users, for all the examined scenarios. Furthermore, we observe that if all the users are considered highly valuable in the crowdsourcing system (High $q_n^*$ GCB scenario), then, the rewards allocation difference among the MAs is balanced, as all the users can contribute high levels of quality and quantity of information (Fig. 5.1b). In contrast, when all the users are less valuable in terms of contributing information (Low $q_n^*$ GCB), then, a greater heterogeneity is observed in the rewards allocation difference (Fig. 5.1b), as the MA with large budget competes aggressively to collect information from the relatively more valuable users (i.e., users with higher ID).

On the other hand, it is interesting to observe that the CCT+GCB scenario, which combines high heterogeneity regarding the users' potential to provide information and at the same time highly valuable users (Fig. 5.1a), results in a similar balanced trend regarding the MAs' rewards allocation difference (Fig. 5.1b) as in the High $q_n^*$ GCB scenario. Moreover, under the information asymmetry case (ICT+GCB), we observe that similar trend is achieved as in the complete information scenario (CCT+GCB), as the MAs realistically derive the users' social characteristics by interacting with them via the personalized contracts introduced in Section 3. Thus, a more balanced trend is derived regarding the users' information contribution (Fig. 5.1a), resulting in similar behavior and values of the MAs' rewards allocation differences, as in the CCT+GCB scenario. Finally, focusing on the MAs'

Table 5.1: Marketing agencies' utilities difference under different budget and information availability scenarios.

| Scen. / Users | GCB | | ICT+GCB | | CCT+GCB | |
|---|---|---|---|---|---|---|
| | S1≡S2 | S3 | S1≡S2 | S3 | S1≡S2 | S3 |
| 1 | 0.000003 | 0.000004 | 0.009853 | 0.010734 | 0.027935 | 0.030443 |
| 2 | 0.000595 | 0.000712 | 0.010820 | 0.011743 | 0.032805 | 0.035521 |
| 3 | 0.002082 | 0.002305 | 0.011783 | 0.012746 | 0.037343 | 0.040240 |
| 4 | 0.005723 | 0.006092 | 0.018697 | 0.019908 | 0.059343 | 0.062987 |
| 5 | 0.008852 | 0.009310 | 0.024662 | 0.026051 | 0.076315 | 0.080442 |
| 6 | 0.010691 | 0.011194 | 0.029050 | 0.030556 | 0.088075 | 0.092504 |
| 7 | 0.011448 | 0.011969 | 0.030066 | 0.031597 | 0.090866 | 0.095363 |
| 8 | 0.015766 | 0.016375 | 0.042645 | 0.044465 | 0.117297 | 0.122398 |
| 9 | 0.021661 | 0.022373 | 0.058134 | 0.060254 | 0.146383 | 0.152074 |
| 10 | 0.028741 | 0.029561 | 0.062900 | 0.065104 | 0.155810 | 0.161679 |

utilities (Fig. 5.1c), we observe that the more valuable are the users and the more information is available to the MAs regarding the users' social characteristics (as in the CCT+GCB case), the higher utility is achieved by the MA with the larger budget. It is reminded that the utility of the MA with the low amount of budget is complementary to the other MA's utility.

## 5.2 Market-driven Analysis

In this section, a detailed market-driven analysis is presented under different realistic scenarios of budget and information availability to the marketing agencies. Initially, three basic scenarios ($S$) are considered: (i) $S_1$: $R^A = 500$, $R^B = 490$, $D = R^A - R^B = 10$; (ii) $S_2$: $R^A = 100$, $R^B = 90$, $D = 10$, and (iii) $S_3$: $R^A = 500$, $R^B = 450$, $D = 50$. For demonstration purposes the following approaches are considered: the GCB with low $q_n^*$ values, CCT+GCB, and ICT+GCB, as presented in the previous section. The results reveal that if the MAs' total budget difference is the same ($S_1$ and $S_2$ scenarios) in terms of absolute value, then, their utilities difference is also the same (Table 5.1), as their actual utilities will be the same (Eq. 2.1), regardless of each individual MAs' total budget availability. Also, the results show that the higher is the total budget difference among the MAs ($S_3$), the higher

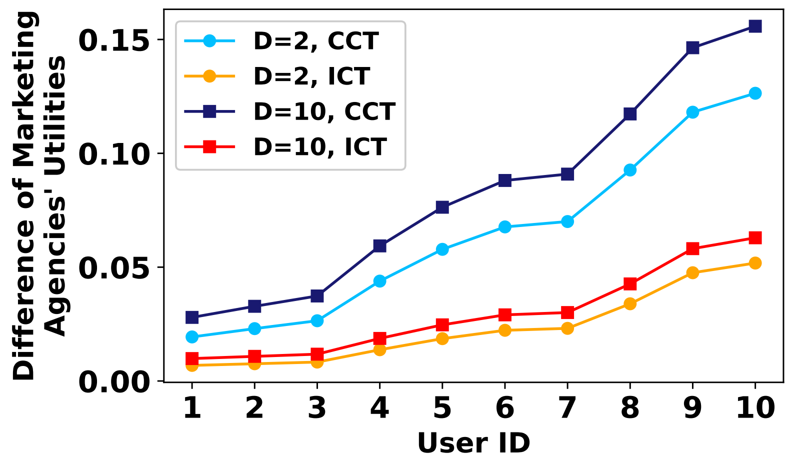


Figure 5.2: Budget difference and information incompleteness.

utilities difference is observed, as the MA with higher budget availability will more aggressively compete for collecting the users' information. Moreover, we observe that the information asymmetry, i.e., the less information is available to the MAs regarding the participating users' social characteristics, favors the MA with lower available budget, as the MA with higher budget looses its privilege and advantage, given the limited information to exploit the users' potential to provide crowdsourced information.

We extend our analysis by considering two scenarios of increasing total budget difference between the MAs, i.e., (i) $R^A = 492$, $R^B = 490$, $D = 2$; (ii) $R^A = 500$, $R^B = 490$, $D = 10$, under the complete (CCT) and incomplete information (ICT) cases. Fig. 5.2 illustrates the MAs' utilities difference as a function of the users' ID under all the examined scenarios. The results confirm our earlier observation that the information asymmetry favors the MA with lower total budget. In particular, we notice that the MAs' utilities difference are lower under the ICT case, as well as the gap between the two scenarios with incomplete information (ICT) decreases compared to the corresponding cases with complete information (CCT) regarding the users' social characteristics.

Finally, we study a scenario, where the total market power is constant, e.g., $R^A + R^B = 200$, under budget asymmetry between the MAs, i.e., $(S_1) : R^A = 101, R^B = 99, D = 2; (S_2) : R^A = 110, R^B = 90, D = 20; (S_3) : R^A = 150, R^B = 50, D = 100$, under the realistic incomplete information case (ICT). Fig. (5.3a) – (5.3b) present the MAs ' rewards allocation and utilities difference, respectively, as a function of the users' ID for all the examined scenarios. The results reveal that as the MAs become more heterogeneous regarding their available budgets, the MA with the larger available budget dominates the crowdsourcing market compared to the one with lower budget by providing higher rewards to the users, and enjoying higher utility. Thus, both the MAs' rewards allocation and utilities differences (Fig. 5.3b) increase for higher $D$ values, and the differences become larger for the more valuable users (i.e., users with higher ID).

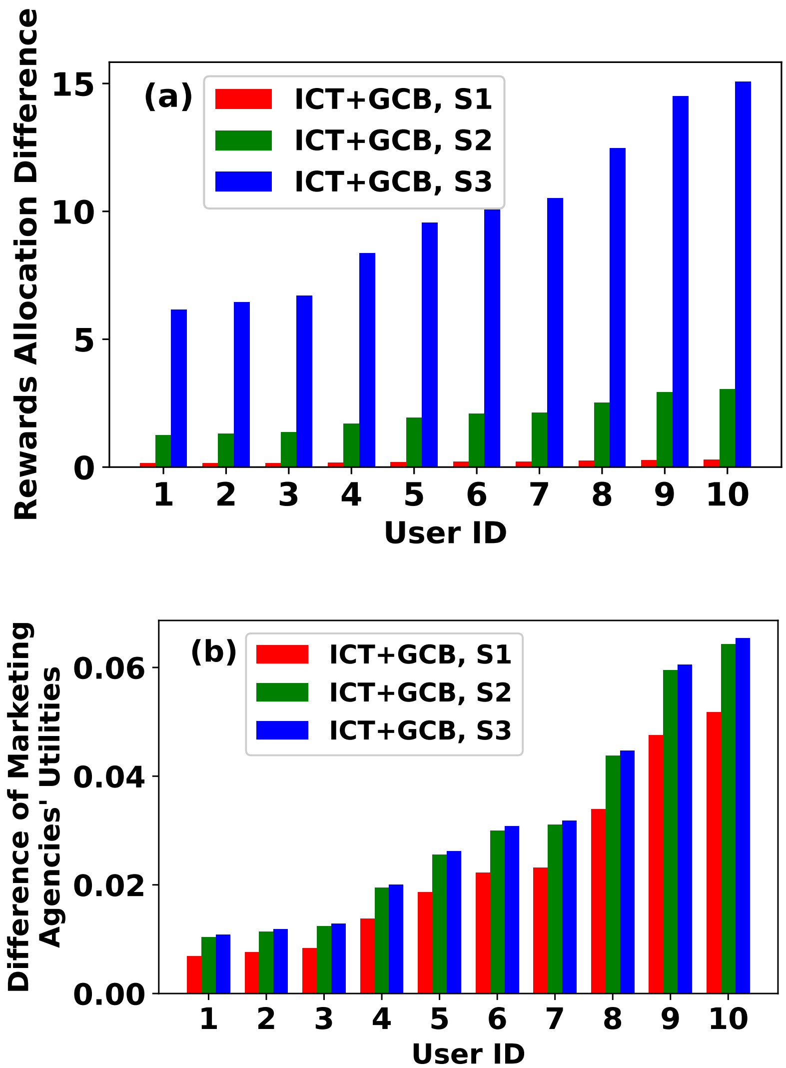


Figure 5.3: Impact of budget asymmetry between the MAs.

# Chapter 6

# Conclusions and Future Works

In this thesis, a network economics-based crowdsourcing framework is introduced in online social networks with competing marketing agencies, based on the principles of labor economics and Colonel Blotto game. Initially, a contract-theoretic framework is introduced to derive the participating users' potential to provide different levels of quality and quantity of information. Then, a Colonel Blotto game is formulated among the marketing agencies to determine the optimal rewards that should be provided to the participating users in order for the latter ones to report their personal information and the marketing agencies to maximize their benefit, i.e., utility, from the crowdsourcing process. A detailed numerical evaluation is presented under different scenarios to reveal the benefits of the proposed crowdsourcing model. Future work for this proposed model would be the inclusion of privacy aware users who are reluctant to reveal their personal data on online social networks. This work would include a variety of different users who all have different levels of risk-awareness when it comes to exposing personal data, which would capture more accurately a users true behavior when participating in crowdsourcing mechanisms.

# References

[1] A. B. Rahman, M. S. Siraj, N. Kubiak, E. E. Tsiropoulou, and S. Papavassiliou, "Network economics-based crowdsourcing in online social networks," in *GLOBECOM 2022-2022 IEEE Global Communications Conference*, pp. 4655–4660, IEEE, 2022.

[2] D. Palacios-Marqués, J. F. Gallego-Nicholls, and M. Guijarro-García, "A recipe for success: Crowdsourcing, online social networks, and their impact on organizational performance," *Technological Forecasting and Social Change*, vol. 165, p. 120566, 2021.

[3] N. Patrizi, S. K. LaTouf, E. E. Tsiropoulou, and S. Papavassiliou, "Museum and visitor interaction & feedback orchestration enabled by labor economics," *IEEE Trans. on Comput. Social Syst.*, pp. 1–12, 2022.

[4] A. S. Al-Nafisa and A. Alnafessah, "Crowdsourcing in digital marketing," in *23rd Int. Conf. on Inf. Integr. & Web Intel.*, pp. 640–644, 2021.

[5] N. Patrizi, E. E. Tsiropoulou, and S. Papavassiliou, "Health data acquisition from wearable devices during a pandemic: A techno-economics approach," in *ICC 2021 - IEEE International Conference on Communications*, pp. 1–6, 2021.

[6] F. Sangoleye, N. Irtija, and E. E. Tsiropoulou, "Data acquisition in social internet of things based on contract theory," in *ICC 2021 - IEEE International Conference on Communications*, pp. 1–6, 2021.

[7] Y. Chen, B. Li, and Q. Zhang, "Incentivizing crowdsourcing systems with network effects," in *IEEE INFOCOM*, pp. 1–9, IEEE, 2016.

[8] Z. Liao, X. Xu, X. Fan, Y. Zhang, and S. Yu, "Grbmc: An effective crowdsourcing recommendation for workers groups," *Expert Systems with Applications*, vol. 179, p. 115039, 2021.

[9] F. Sangoleye, M. S. Hossain, E. E. Tsiropoulou, and J. Plusquellic, "Network economics-based crowdsourcing in uav-assisted smart cities environments," in *2022 18th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 101–108, 2022.

[10] S. Xu, L. Liu, L. Cui, Q. Li, and Z. Yan, "Promoting higher revenues for both crowdsourcer and crowds in crowdsourcing via contest," in *2019 IEEE International Conference on Web Services (ICWS)*, pp. 403–407, 2019.

[11] J. Xu, C. Guan, H. Wu, D. Yang, L. Xu, and T. Li, "Online incentive mechanism for mobile crowdsourcing based on two-tiered social crowdsourcing architecture," in *2018 15th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON)*, pp. 1–9, 2018.

[12] Y. Liu, Y. Xu, and S. Qin, "What are key components when creating an innovative crowdsourcing business model," in *2018 24th International Conference on Automation and Computing (ICAC)*, pp. 1–4, 2018.

[13] W. Xiulan and W. Shiyan, "An optimal assignment for mobile sensing tasks in spatial crowdsourcing," in *2016 5th International Conference on Computer Science and Network Technology (ICCSNT)*, pp. 681–687, 2016.

[14] A. R. Kurup and G. P. Sajeev, "Task personalization for inexpertise workers in incentive based crowdsourcing platforms," in *2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 286–292, 2018.

[15] E. E. Tsiropoulou, A. Thanou, S. T. Paruchuri, and S. Papavassiliou, "Self-organizing museum visitor communities: A participatory action research based approach," in *2017 12th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, pp. 101–105, IEEE, 2017.

[16] M. Brambilla and A. Bozzon, "Web data management through crowdsourcing upon social networks," in *2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 1123–1127, 2012.

[17] E. E. Tsiropoulou, A. Thanou, and S. Papavassiliou, "Quality of experience-based museum touring: A human in the loop approach," *Social Network Analysis and Mining*, vol. 7, pp. 1–13, 2017.

[18] A. Thanou, E. E. Tsiropoulou, and S. Papavassiliou, "Quality of experience under a prospect theoretic perspective: A cultural heritage space use case," *IEEE Transactions on Computational Social Systems*, vol. 6, no. 1, pp. 135–148, 2019.

[19] C. Huang, H. Yu, J. Huang, and R. A. Berry, "Crowdsourcing with heterogeneous workers in social networks," in *IEEE GLOBECOM*, pp. 1–6, IEEE, 2019.

[20] W. Wang, Z. He, P. Shi, W. Wu, Y. Jiang, B. An, Z. Hao, and B. Chen, "Strategic social team crowdsourcing: Forming a team of truthful workers for crowdsourcing in social networks," *IEEE Transactions on Mobile Computing*, vol. 18, no. 6, pp. 1419–1432, 2018.

[21] Y. Xu, M. Xiao, J. Wu, S. Zhang, and G. Gao, "Incentive mechanism for spatial crowdsourcing with unknown social-aware workers: A three-stage stackelberg game approach," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2022.

[22] Z. Saberi, O. Hussain, M. Saberi, and E. Chang, "Stackelberg game-theoretic approach in joint pricing and assortment optimizing for small-scale online retailers: Seller-buyer supply chain case," in *2018 IEEE 32nd International Conference on Advanced Information Networking and Applications (AINA)*, pp. 834–838, 2018.

[23] D. Z. Tootaghaj and F. Farhat, "Cage: A contention-aware game-theoretic model for heterogeneous resource assignment," in *2017 IEEE International Conference on Computer Design (ICCD)*, pp. 161–164, 2017.

[24] M. S. Siraj, M. S. Hossain, R. Brown, E. E. Tsiropoulou, and S. Papavassiliou, "Incentives to learn: A location-based federated learning model," in *2022 Global Information Infrastructure and Networking Symposium (GIIS)*, pp. 40–45, 2022.

[25] V. Leon and S. R. Etesami, "Bandit learning for dynamic colonel blotto game with a budget constraint," in *2021 60th IEEE Conference on Decision and Control (CDC)*, pp. 3818–3823, 2021.

[26] E. E. Tsiropoulou, G. Kousis, A. Thanou, I. Lykourentzou, and S. Papavassiliou, "Quality of experience in cyber-physical social systems based on reinforcement learning and game theory," *Future Internet*, vol. 10, no. 11, p. 108, 2018.

[27] C. A. Kamhoua, K. A. Kwiat, and J. S. Park, "A game theoretic approach for modeling optimal data sharing on online social networks," in *2012 9th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, pp. 1–6, 2012.

[28] C. Griffin and A. Squicciarini, "Toward a game theoretic model of information release in social media with experimental results," in *2012 IEEE Symposium on Security and Privacy Workshops*, pp. 113–116, 2012.

[29] F. Zhan, G. Laines, S. Deniz, S. Paliskara, I. Ochoa, I. Guerra, S. Tayeb, C. Chiu, M. Pirouz, E. Ploutz, J. Zhan, L. Gewali, and P. Oh, "Prediction of online social networks users' behaviors with a game theoretic approach," in *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, pp. 1–2, 2018.

[30] A. Castiglione, G. Cozzolino, V. Moscato, and G. Sperlì, "Analysis of community in social networks based on game theory," in *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, pp. 619–626, 2019.

[31] Y. Li, Y. Li, H. Hu, H. V. Zhao, and Y. Chen, "Graphical evolutionary game theoretic analysis of super users in information diffusion," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5650–5654, 2020.

[32] X. Cao, Y. Chen, and K. J. R. Liu, "Understanding popularity dynamics: Decision-making with long-term incentives," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 3, no. 1, pp. 91–103, 2017.

[33] E. E. Tsiropoulou, A. Thanou, and S. Papavassiliou, "Modelling museum visitors' quality of experience," in *2016 11th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*, pp. 77–82, IEEE, 2016.

[34] E. Stai, S. Kafetzoglou, E. E. Tsiropoulou, and S. Papavassiliou, "A holistic approach for personalization, relevance feedback & recommendation in enriched multimedia content," *Multimedia Tools and Applications*, vol. 77, pp. 283–326, 2018.

[35] V. Pouli, S. Kafetzoglou, E. E. Tsiropoulou, A. Dimitriou, and S. Papavassiliou, "Personalized multimedia content retrieval through relevance feedback techniques for enhanced user experience," in *2015 13th International Conference on Telecommunications (ConTEL)*, pp. 1–8, 2015.

[36] L. Wang, Z. Yu, D. Yang, T. Wang, E. Wang, B. Guo, and D. Zhang, "Task execution quality maximization for mobile crowdsourcing in geo-social networks," *Proceedings of the ACM on Human-Computer Interaction*, vol. 5, no. CSCW2, pp. 1–29, 2021.

[37] K. Ntalianis and N. Tsapatsoulis, "Wall-content selection in social media: A revelance feedback scheme based on explicit crowdsourcing," in *2016 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, pp. 534–539, 2016.

[38] L. Wang, D. Yang, Z. Yu, Q. Han, E. Wang, K. Zhou, and B. Guo, "Acceptance-aware mobile crowdsourcing worker recruitment in social networks," *IEEE Transactions on Mobile Computing*, 2021.

[39] P. A. Apostolopoulos, E. E. Tsiropoulou, and S. Papavassiliou, "Risk-aware social cloud computing based on serverless computing model," in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, 2019.

[40] Z. Wang, Y. Huang, X. Wang, J. Ren, Q. Wang, and L. Wu, "Socialrecruiter: Dynamic incentive mechanism for mobile crowdsourcing worker recruitment with social networks," *IEEE Transactions on Mobile Computing*, vol. 20, no. 5, pp. 2055–2066, 2020.

[41] H.-H. Shuai, C.-Y. Shen, H.-C. Hsu, D.-N. Yang, C.-K. Chou, J.-H. Lin, and M.-S. Chen, "Revenue maximization for telecommunications company with social viral marketing," in *2015 IEEE International Conference on Big Data (Big Data)*, pp. 1306–1310, 2015.

[42] Y. Wang, W. Dai, Q. Jin, and J. Ma, "Bcinet: A biased contest-based crowdsourcing incentive mechanism through exploiting social networks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 8, pp. 2926–2937, 2020.

[43] L. Hao, C. Jin, X. Gao, F. Wu, and G. Chen, "Towards cost-effective and budget-balanced task allocation in crowdsourcing systems," in *2017 IEEE 36th International Performance Computing and Communications Conference (IPCCC)*, pp. 1–8, 2017.

[44] S. Aghababaei, E. Gultepe, I. Chepurna, and M. Makrehchi, "Activity-based sampling of twitter users for temporal prediction models," in *2016 International Conference on Behavioral, Economic and Socio-cultural Computing (BESC)*, pp. 1–6, 2016.

[45] B.-W. Chen and W. Ji, "Intelligent marketing in smart cities: Crowdsourced data for geo-conquesting," *IT Professional*, vol. 18, no. 4, pp. 18–24, 2016.

[46] E. Mao and J. Zhang, "What drives consumers to click on social media ads? the roles of content, media, and individual factors," in *2015 48th Hawaii International Conference on System Sciences*, pp. 3405–3413, 2015.

[47] V. Boonjing and D. Pimchangthong, "Data mining for customers' positive reaction to advertising in social media," in *2017 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pp. 945–948, 2017.

[48] R. Machedon, W. Rand, and Y. Joshi, "Automatic crowdsourcing-based classification of marketing messaging on twitter," in *2013 International Conference on Social Computing*, pp. 975–978, 2013.

[49] S. Wibowo, R. Hidayat, Y. Suryana, D. Sari, and U. Kaltum, "Measuring the effect of advertising value and brand awareness on purchase intention through the flow experience method on facebook's social media marketing big data," in *2020 8th International Conference on Cyber and IT Service Management (CITSM)*, pp. 1–5, 2020.

[50] H. De Silva, P. Jayasinghe, A. Perera, S. Pramudith, and D. Kasthurirathna, "Social media based personalized advertisement engine," in *2017 11th International Conference on Software, Knowledge, Information Management and Applications (SKIMA)*, pp. 1–6, 2017.

[51] R. Jiamthapthaksin and T. H. Aung, "User preferences profiling based on user behaviors on facebook page categories," in *2017 9th International Conference on Knowledge and Smart Technology (KST)*, pp. 248–253, 2017.

[52] Y. Ouyang, L. Huang, and Z. Wei, "Research on the influence mechanism of precision marketing on online consumer behaviour based on the sor paradigm and eroglus (2001) online shopping response model : Take the tiktok platform as an example," in *2021 2nd International Conference on E-Commerce and Internet Technology (ECIT)*, pp. 218–223, 2021.

[53] A. Gogulamudi and S. Prabhu, "Consumer decisions under the influence of social media & behavioral targeting," in *2022 International Conference on Decision Aid Sciences and Applications (DASA)*, pp. 357–361, 2022.

[54] W. Peng, "New media e-commerce marketing investment algorithm based on user preferences data mining," in *2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, pp. 976–979, 2021.

[55] C. K. Coursaris, W. Van Osch, and B. A. Balogh, "Do facebook likes lead to shares or sales? exploring the empirical links between social media content,

brand equity, purchase intention, and engagement," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pp. 3546–3555, 2016.

[56] H. N. Bhor, T. Koul, R. Malviya, and K. Mundra, "Digital media marketing using trend analysis on social media," in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, pp. 1398–1400, 2018.

[57] F. Kaske, M. Kugler, and S. Smolnik, "Return on investment in social media–does the hype pay off? towards an assessment of the profitability of social media in organizations," in *2012 45th Hawaii International Conference on System Sciences*, pp. 3898–3907, 2012.

[58] M. Diamanti, G. Fragkos, E. E. Tsiropoulou, and S. Papavassiliou, "Unified user association and contract-theoretic resource orchestration in noma heterogeneous wireless networks," *IEEE Open Journal of the Communications Society*, vol. 1, pp. 1485–1502, 2020.

[59] M. Diamanti, P. Charatsaris, E. E. Tsiropoulou, and S. Papavassiliou, "Incentive mechanism and resource allocation for edge-fog networks driven by multi-dimensional contract and game theories," *IEEE Open Journal of the Communications Society*, vol. 3, pp. 435–452, 2022.

[60] L. Sigerson and C. Cheng, "Scales for measuring user engagement with social network sites: A systematic review of psychometric properties," *Computers in Human Behavior*, vol. 83, pp. 87–105, 2018.

[61] A. Thanou, E. E. Tsiropoulou, and S. Papavassiliou, "A sociotechnical approach to the museum congestion management problem," *IEEE Transactions on Computational Social Systems*, vol. 7, no. 2, pp. 563–568, 2020.

[62] P. Bolton and M. Dewatripont, *Contract theory*. MIT press, 2004.

[63] N. Patrizi, G. Fragkos, E. E. Tsiropoulou, and S. Papavassiliou, "Contract-theoretic resource control in wireless powered communication public safety systems," in *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, pp. 1–6, 2020.

[64] N. Irtija, F. Sangoleye, and E. E. Tsiropoulou, "Contract-theoretic demand response management in smart grid systems," *IEEE Access*, vol. 8, pp. 184976–184987, 2020.

[65] N. Patrizi, S. K. LaTouf, E. E. Tsiropoulou, and S. Papavassiliou, "Prosumer-centric self-sustained smart grid systems," *IEEE Systems Journal*, vol. 16, no. 4, pp. 6042–6053, 2022.

[66] B. Roberson, “The colonel blotto game,” *Economic Theory*, vol. 29, no. 1, pp. 1–24, 2006.

[67] A. Ferdowsi, A. Sanjab, W. Saad, and T. Basar, “Generalized colonel blotto game,” in *Annual Am. Control Conf.*, pp. 5744–5749, IEEE, 2018.