Spring 5-14-2022

Decentralized Intelligent Decision Making in Cyber Physical Social Systems

Nathan Patrizi

Decentralized Intelligent Decision Making in Cyber Physical Social Systems

by

Nathan Patrizi

B.S, University of New Mexico, 2019
Master of Science, University of New Mexico, 2020

PH.D. DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Ph.D.
Computer Engineering

The University of New Mexico

Albuquerque, New Mexico

February, 2022

Dedication

To my parents, who have always supported me and helped me throughout my life.

Acknowledgments

I would first like to thank my advisor, Professor Eirini Eleni Tsiropoulou, who has helped me invaluably throughout my graduate and undergraduate studies, by providing advice when needed and pushing me to succeed.

I would also like to thank Professor Jim Plusquellic for being a part of my dissertation committee along with providing excellent courses in both my graduate and undergraduate degrees. Covering interesting and detailed topics, relating to hardware-oriented security and design, with real world applications.

I would like to thank Professor Payman Zarkesh-Ha for being a member of my dissertation committee.

I would like to thank Professor Symeon Papavassiliou for being a member of my dissertation committee and collaboration throughout my graduate studies.

Lastly, I would like to thank Pavlos Athanasios Apostolopoulos and George Fragkos for their help in performing the research presented in this Proposal.

Decentralized Intelligent Decision Making in Cyber Physical Social Systems

by

Nathan Patrizi

B.S, University of New Mexico, 2019

Master of Science, University of New Mexico, 2020

Ph.D. Proposal, University of New Mexico, 2022

Abstract

The accelerated evolution towards jointly considering the physical, cyber, and social space is expected to dramatically increase the interest of the research and industrial community to build efficient, resilient, and secure Cyber Physical Social Systems. In this dissertation, we focus our research activities on devising decentralized intelligent decision making models, frameworks, and algorithms to support the smooth operation of Cyber Physical Social Systems. The proposed decentralized intelligent decision making models are jointly exploiting theories from the field of Economics, such as Game Theory and Contract Theory, and from the field of Computer Science, such as Reinforcement Learning concepts. Reinforcement learning is applied to allow for humans to make informed decisions in the considered Cyber Physical Social Systems based off of the dynamically changing environment around them. Additionally, contract theoretic and game theoretic models allow for us to accurately depict the relationships between the different involved entities in the examined system. Several research problems have been examined which can be sum-

marized as follows: (i) socio-physical human orchestration in smart cities, (ii) socio-aware public safety framework design, (iii) unmanned aerial vehicle or UAV-enabled dynamic multi-target tracking and sensing framework, (iv) resource orchestration in wireless powered communication public safety systems, (v) health data acquisition from wearable devices during a pandemic by following a techno-economics approach, (vi) museum and visitor interaction and feedback orchestration enabled by labor economics, and (vii) design and operation of prosumer-centric self-sustained smart grid systems. Finally, all the above problems are thoroughly evaluated and tested via a series of simulations and emulations with regards to the main characteristics of their operation, as well as against other approaches from the literature.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

Many recent advances in cyber-physical systems have allowed for new and more sophisticated interactions in smart cities [1]. As there is more access to information and data, it becomes more important to not just utilize or provide data, but to do so in an intelligent and informed manner [2]. Providing this data and utilizing it correctly helps facilitate better smart services throughout people's lives as these systems will have the information available to them to make informed decisions about how to best provide for their individuals [3]. Additionally, it is important to more accurately model human decision making, while it would be easy to simply provide all information to these smart systems, it would not accurately model human decision making [4]. Human's will naturally exhibit risk averse decision making behaviours, a reluctance to participate without a reward, and competitiveness amongst each other while sharing the same pool of resources [5].

Motivated by the above observations and due to the rising human-centric technological achievements, Cyber Physical Social Systems (CPSS) arise as a new Cyber

Physical Systems paradigm that encompasses the digital fusion among human, computers, networks, and smart objects and devices [6]. CPSS consist not only of raw sensing and actuating hardware and software, but also consider the humans behavior, actions, decisions, interactions, and social characteristics in order to plan their efficient operation [7]. Given the joint consideration of all those heterogeneous aspects within CPSS, these systems are characterized as complex and nowadays, they lack effective design approaches to systematically study their operation [8]. The ultimate goal of the CPSS is to bridge the gap among the Cyber Physical and the Cyber Social systems in order to meet the humans' social interaction demands and appropriately adapt to the physical world conditions. Indicative applications of the CPSS include, but are not limited to, smart home [9], smart cities [10, 11], autonomous vehicles [12], recommendation and advertisement systems [13, 14, 15], smart medical services [16, 17], smart grid systems [18, 19, 20], smart agriculture [21, 22, 23], public safety [24, 25], interactive cultural spaces [26], smart secure systems [27, 28, 29], just to name a few of them [30, 31].

CPSS is developed in a virtual three dimensional environment, consisting of the (i) social, (ii) physical, and (iii) cyberspace dimensions. The building components of the CPSS are (1) the sensing devices, (2) physical objects and smart things [32], (3) the humans, and the (4) networking [33], computing, communications, and control functionalities [34, 35, 36]. Within the created three dimensional research space of the CPSS, the main research challenges that have been identified are listed below [37, 38]:

1. Human behavior and interaction with the environment
2. Human-computer interaction
3. Context awareness and management
4. Device management and discovery

5. Social computing

6. Seamless migration technologies

7. Security, and privacy

Given the increased heterogeneity and complexity of the CPSS, the methodologies, which should be devised in order to deal with the aforementioned research challenges, should be of distributed and decentralized nature. The latter observation is a fundamental principle within CPSS, as centralized approaches could not scale in such complex systems for the following indicative reasons: (a) the physical infrastructure is owned by different service, internet, network, computing, content providers, (b) the humans social and behavioral characteristics are private and known only by each entity itself, (c) the cyberspace is by nature distributed with out having a single point of control. The decentralized nature of the CPSS motivates the study of decentralized approaches that can provide the enhanced flexibility both to the CPSS as a whole, and to its entities to act in an autonomous and intelligent manner and adapt in a real-time manner in the dynamic changes of the environment.

Based on the above observation, this Ph.D. dissertation is motivated by the need of devising decentralized methodologies, frameworks, models, and algorithms in order to study the CPSS and their variety of functionalities, applications, and services that they can offer to the end-users, i.e., the humans. In the following subsection, we present the main characteristics and properties of three main theoretical techniques that have been used in order to perform the presented research. Those techniques are: (i) Contract Theory [39], (ii) Game Theory [40], and (iii) Reinforcement Learning [41]. Based on those techniques, which define the fundamentals of this Ph.D. dissertation, several applications within smart cities scenarios have been examined, such as human orchestration in visiting points of interest [42], data collection from citizens in public safety scenarios [43], multi-target tracking and sensing assisted by

Unmanned Aerial Vehicles [44], resource management in wireless powered communications [45], health data acquisition during a pandemic, humans interactions within cultural places, and demand response management within smart grid systems.

1.2 Decentralized Intelligent Decision-Making

1.2.1 Contract-theoretic Models

Contract theory models the relationship and interactions between an employer and an employee [39]. This model allows for expectations for effort and reward to be balanced based on the capability of each employee, with the employee providing an appropriate amount of effort based on their own individual ability [46]. This results in the employer providing a personalized reward that encourages getting this effort from the employee.

Specifically, Contract Theory (CT), lying in the area of Labor Economics, provides the mathematical foundations to create mutually agreeable contracts or arrangements between economic players, i.e., principal or employer(s) and agents or employees, in presence of complete or incomplete information (often referred to as asymmetric information). The incompleteness of information refers to the unknown by the principal agents' private characteristics that under typical circumstances steer the contract formulation. Under this concept, the principal creates contract bundles based on the statistical knowledge of the potential agents' private information, i.e., the agents' types, to motivate them provide back their effort and hence, reveal their actual type [47].

As a means of providing the appropriate incentives to the humans to cooperate in the direction of the CPSS's ultimate objective, Contract Theory encompasses the notion of the end-users' personal utility satisfaction, i.e., the end-users' achievement of a payoff greater or equal to a threshold value. Therefore, the maximization of the ultimate CPSS's objective is pursued subject to the end-users' personal utility functions satisfaction [48]. From the CPSS's perspective, the principal and the agents can correspond to different entities of the CPSS architecture under consideration and can target different metrics. Under this concept, a wide variety of optimization problems can be formulated, concurrently, targeting different metrics from the system's and the humans' perspective.

In the following subsections, we present the two major models of agency problems that are formulated and solved under the principles of Contract Theory, namely the Adverse Selection and the Moral Hazard, whose potential of expressing typical CPSS problems (e.g., crowd-sourcing [49], human-orchestration under different principal and agents' utility functions) will be evaluated in the context of the Ph.D. dissertation.

1.2.2 Adverse Selection Problem

One of the most common problems that arises between a principal and an agent that falls into the range of adverse selection problem modeling is the "employment" contract, under which the agent's desired performance/effort by the principal and the principal's reward back to the agent, are agreed. Specifically, the principal is unaware of the prospective agent's capabilities, i.e., the agent's private information, and tries to elicit this private information via its contract offer. Following the revelation principle, the principal can offer multiple employment contracts destined to different-capability agents, and each agent selects the appropriate contract offer for its type, i.e., the one that maximizes its personal utility. As such, the agent ultimately reveals its actual type to the principal.

Let us consider that there are N different agent types, denoted as θ_i, $i \in \{1, \cdots, N\}$

that bear different private information. Although there exists information asymmetry between the principal and the agents, the principal possesses statistical information regarding the existence/occurrence of different agent types. Hence, we define as λ_i the probability of facing the agent type θ_i, such that $\sum_{i=1}^{N} \lambda_i = 1$. The contract bundle designed and offered by the principal to each agent i is denoted as $\{p_i, r_i\}$, where p_i corresponds to the agent's effort wanted by the principal and r_i is the principal's reward provided back to the agent. Therefore, we formulate the principal's expected utility function as the principal's expected profit by the agents' efforts minus their provided rewards, i.e., $U_{pr} = \sum_{i=1}^{N} [\lambda_i \cdot (p_i - \mathcal{C} \cdot r_i)]$, where $\mathcal{C} \in \mathbb{R}^+$ is the principal's unit cost of its provided reward to each agent. In a similar manner, the agent's i personal utility function is defined as $U_i = \theta_i \cdot e(r_i) - p_i$, where the first term expresses the agent's evaluation of its received reward minus its provided effort. Specifically, the agent's evaluation function of reward $e(r_i)$ is strictly increasing and concave with respect to the agent's i received reward (i.e., $e(0) = 0$, $e'(r_i) > 0$, $e''(r_i) < 0$) and is commonly modeled as $\sqrt{r_i}$ or $\log(1 + r_i)$.

Following the adverse selection problem formulation, the principal's utility function U_{pr} is maximized subject to the agents' satisfaction of their personal utilities U_i, $\forall i \in \{1, \ldots, N\}$, expressed by the Individual Rationality and Incentive Compatibility constraints, as described below.

Definition 1. ***(Individual Rationality (IR))*** *A contract bundle $\{p_i, r_i\}$ satisfies the individual rationality constraint if each agent receives a non-negative utility, i.e.,*

$$\theta_i \cdot e(r_i) - p_i \geq 0, \forall i \in \{1, \ldots, N\}. \tag{1.1}$$

Definition 2. ***(Incentive Compatibility (IC))*** *Each agent must select the contract bundle $\{p_i, r_i\}$ that is designed specifically for their own type θ_i, i.e.,*

$$\theta_i \cdot e(r_i) - p_i \geq \theta_i \cdot e(r_i') - p_i', \forall i, i' \in \{1, \ldots, N\}, i \neq i'. \tag{1.2}$$

The IR constraint ensures the participation of each agent in the contract agreement, by marginally satisfying the agent's personal utility function, while the IC constraint guarantees that each agent can only receive the highest utility when selecting the contract bundle designed for its own type. Therefore, the optimization problem to be solved can be written as,

$$\max_{\{p_i,r_i\}_{\forall i\in\{1,\dots,N\}}} U_{pr} = \sum_{i=1}^{N}[\lambda_i \cdot (p_i - \mathcal{C} \cdot r_i)] \tag{1.3a}$$

$$\textbf{s.t. } \theta_i \cdot e(r_i) - p_i \geq 0, \forall i \in \{1,\dots,N\} \tag{1.3b}$$

$$\theta_i \cdot e(r_i) - p_i \geq \theta_i \cdot e(r_{i'}) - p_{i'}, \forall i, i' \in \{1,\dots,N\}, i \neq i' \tag{1.3c}$$

It should be noted that the adverse selection problem model presented in the current section corresponds to the discrete agent type case and can, also, be generalized to the continuous agent type case to fit more realistic scenarios.

1.2.3 Moral Hazard Problem

In the adverse selection problem formulation, the notions of agent's effort and agent's performance were interchangeably used, assuming that a specific amount of effort yields in a proportional manner to an amount of performance. However, in several realistic scenarios, the agent's effort is costly and its ultimate performance observed by the principal differs from the effort that has been actually exerted. In order to model such problems, where the agent's effort is hidden and only the final performance is observable by the principal, the moral hazard problem formulation is adopted.

According to the basic moral hazard model, the agent's performance q is defined as a noisy signal of its actual provided effort a, such as $q = a + \varepsilon_q$, where $\varepsilon_q \sim N(\mu_q, \sigma_q^2)$. Given that the principal is unaware of the agent's effort, the principal has to strategically reward the agent considering a double compensation scheme that includes a fixed reward t and a variable s. The fixed amount of reward is intended to incentivize the agent to provide its best effort and hence, is offered while signing the contract. On the contrary, the variable reward is offered as long as the principal observes the agent's ultimate performance and its purpose is to compensate the agent's incurred cost of providing its best effort. Thus, the total reward provided to the agent is defined as $r = t + s \cdot q$. The agent is assumed to have constant absolute risk averse (CARA) preferences, meaning that the agent's attitude towards risk is constant as its reward increases. As a result, we formulate the agent's personal utility as $U_a = -e^{-\eta[r-\psi(a)]}$, where $\eta > 0$ is the agent's coefficient of absolute risk aversion ($\eta = -U_a''/U_a'$), the higher the value of which corresponds to less incentives for the agent to exert an effort. Also, the term $\psi(a)$ corresponds to the agent's cost function of providing its effort and is assumed to be quadratic, such as $\psi(a) = \frac{1}{2}ca^2$. The principal's utility function is modeled as the evaluation of the agent's ultimate performance minus its total offered compensation, i.e., $U_{pr} = q - r = (1 - s) \cdot a - t$.

Considering the problem description above, the contract bundle designed and offered by the principal to the agent is denoted as $\{a, r\}$, where a corresponds to the agent's actual effort and r is the principal's total provided reward. Similarly to the adverse selection model, the principal's utility U_{pr} is maximized subject to the agent's satisfaction of its personal utility U_a. Thus, the optimization problem to be solved can be written as follows.

$$\max U_{pr} = (1 - s) \cdot a - t \tag{1.4a}$$

$$\textbf{s.t. } E[-e^{-\eta[r-\psi(a)]}] \geq U_{min} \tag{1.4b}$$

$$a \in \underset{a}{\operatorname{argmax}}\, E[-e^{-\eta[r-\psi(a)]}] \tag{1.4c}$$

where U_{min} is the minimum acceptable utility for the agent to sign the contractual agreement. In accordance with the adverse selection model, the principal has to reassure the agent's marginal participation in the contract by satisfying its personal utility function, as imposed by the first constraint of the optimization problem, i.e., the IR constraint. The second constraint maps to the IC constraint and guarantees that the agent can maximize its personal utility when selecting the right amount of effort.

1.2.4 Game Theory & Reinforcement Learning

Game Theory has been recognized as a field of applied mathematics targeting at the study of strategic decision making in conditions of competition, collaboration, and/or conflict, holding its foundations from the book of John von Neumann and Oskar Morgenstern in “Theory of Games and Economic Behavior”, while it was further extended and formalized by John Nash, who mainly focused his research work on non-cooperative games. Game Theory, was initially introduced as a theory related to social and economic disciplines, however, nowadays it has been widely accepted and adopted as a fundamental, useful and powerful tool across various areas including computer engineering, computer science, Internet of Things, Cyber Physical Systems, Cyber Physical Social Systems, business, and wireless networking, among others [50].

Game Theory is built upon the concept of a game, representing an interaction between different rational entities, or players, whose individual decisions affect each other’s payoff and actions, who aim to maximize their expected benefit based on their current status of information. The games are modeled based on the players’ possible and feasible strategies, representing the set of available options to the involved entities under which they define their most beneficial decisions and are determined as pure (if the decision environment is deterministic), or mixed if multiple options can be probabilistically selected during the game. The decisions of the entities lead to a corresponding outcome which provides a payoff or utility to the entities, representing a quantification of the entities gains or losses from their corresponding actions.

Different variations of games exist referring to different conditions of interactions among the entities, i.e., players based on the considered situation.

- Static and dynamic games: The first type of games refers to situations where the involved entities have a certain amount of knowledge which remains the same during the game, while dynamic games imply that users can gain information from their previous actions.

- Zero-sum and non zero-sum games: The first indicates a strictly competitive situation where the benefit of one entity leads to an equivalent loss of the other entities, while the second category refers to a situation where the cumulative gains and losses of the entities are not complementary. The involved entities can be either competitive or non-competitive among each other.

- Non stochastic and stochastic games: In stochastic games, the game is played in stages with a certain state to evolve according to a probabilistic rule.

- Games with complete and incomplete information: Complete information games consider that the information available among the involved entities is common knowledge to everybody involved int he game, while incomplete information games deal with situations where the involved entities are aware of only partial information of the game's characteristics.

- Non-cooperative and cooperative games: In cooperative games the involved entities can form collaborations towards achieving optimal outcomes, while in non-cooperative games the involved entities compete with each other often having conflicting interests [51, 52, 53].

- Games with perfect or imperfect background knowledge: In the first type, the involved entities are fully aware of the history of the game, while in the latter this does not hold true.

Integrating Reinforcement Learning (RL) algorithms and techniques can facilitate the interaction of humans with the CPSS in order to handle in real time aspects like crowdsourcing, crowdsensing, coordination among them, navigation, and many more. Reinforcement Learning can support the humans to learn from their environment and adjust the resource allocation of the system's limited available resources to the needs of its human by introducing intelligent resource management and decision making tools related to the main goal of the CPSS, e.g., public safety, health monitoring, etc. Furthermore, Reinforcement Learning methodologies (e.g., Gradient Ascent Learning, Log Linear Learning, Q-learning) can be applied to different scenarios and can play an important role for the CPSS to operate more autonomously and automatically optimize many of its functions. In such a way, the CPSS can organize itself and lead to decentralized structures, with the humans to become part of the decision making process without pushing large volumes of data to the centralized entities, e.g., Cloud, developing a parallel processing capacity which accelerates the communication and minimizes unnecessary coordination among the humans. It should be noted that a centralized approach would not scale within a CPSS system due to its increased heterogeneity, the dynamic manner that it evolves over time, the plethora of diverse involved entities, and the threat of single point of failure.

Subsequently, the CPSS can optimize its resources by better understanding the priorities of its involved entities, decrease the decision making time, and manage congestion in order to deliver superior quality of services and an overall more holistic experience. The above are already a reality in recent CPSS paradigms (mobile edge caching, mobile edge computing etc.), hence it is of vital importance to invest more in designing Reinforcement Learning techniques into CPSS operations within realistic smart applications, such as smart cities, smart health, and others. In this Ph.D. dissertation, we have elaborated more on the Gradient Ascent Reinforcement Learning algorithms and their applicability in real time decision making in CPSS. The introduced algorithms are coupled either with contract-theoretic or game-theoretic approaches, which jointly enable the involved entities to perform the real decision making by sensing and accounting for the dynamic change of the surrounding environment. The main parameters that have been used in order to evaluate the performance of these algorithms are their execution time and the efficiency of the solutions that they converge based on the main goal of the CPSS at each of the examined research problems.

1.3 Contributions, Publications, and Organization

1.3.1 Contributions

The accelerated evolution towards jointly considering the physical, cyber, and social space is expected to dramatically increased the interest of the research and industrial community to build efficient, resilient, and secure Cyber Physical Social Systems. In this dissertation, we focus our research activities on devising decentralized intelligent decision making models, frameworks, and algorithms to support the smooth operation of Cyber Physical Social Systems. The proposed decentralized intelligent decision making models are jointly exploiting theories from the field of Economics, such as Game Theory and Contract Theory, and from the field of Computer Science, such as Reinforcement Learning concepts. Reinforcement learning is applied to allow for humans to make informed decisions in the considered Cyber Physical Social Systems based off of the environment around them. Additionally, contract theoretic and game theoretic models allow for us to accurately depict the relationships between the different involved entities in the examined system. Several research problems have been examined which can be summarized as follows: (i) socio-physical human orchestration in smart cities, (ii) socio-aware public safety framework design, (iii) unmanned aerial vehicle or UAV-enabled dynamic multi-target tracking and sensing framework, (iv) resource orchestration in wireless powered communication public safety systems, and (v) health data acquisition from wearable devices during a pandemic by following a techno-economics approach, (vi) museum and visitor interaction and feedback orchestration, (vii) prosumer-centric self-sustained smart grid systems. Finally, all the above problems are thoroughly evaluated and tested via a series of simulations and emulations with regards to the main characteristics of their operation, as well as against other approaches from the literature.

1.3.2 Publications

All of the research work presented in this Ph.D. thesis is published or submitted for publication in peer-reviewed journals and conferences. At the beginning of the following list, we present the peer-reviewed and published research papers, while at the end of the list we present separately the research work being submitted and under review.

1. **N. Patrizi**, E.E. Tsiropoulou, and S. Papavassiliou, "Health Data Acquisition from Wearable Devices during a Pandemic: A Techno-Economics Approach," in IEEE ICC, 2021. pp. 1-6, 2021

2. **N. Patrizi**, G. Fragkos, E.E. Tsiropoulou, and S. Papavassiliou, "Contract - Theoretic Resource Control in Wireless Powered Communication Public Safety Systems," in IEEE GLOBECOM, pp. 1-6, 2020.

3. **N. Patrizi**, G. Fragkos, K. Ortiz, M. Oishi, and E.E. Tsiropoulou, "A UAV-enabled Dynamic Multi-Target Tracking and Sensing Framework," in IEEE GLOBECOM, pp. 1-6, 2020.

4. G. Fragkos, **N. Patrizi**, E. E. Tsiropoulou, and S. Papavassiliou, ”Socio-aware Public Safety Framework Design: A Contract Theory based Approach,” ICC 2020 - 2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, pp. 1-7, 2020.

5. **N. Patrizi**, P.A. Apostolopoulos, K. Rael, and E.E. Tsiropoulou, ”Socio-physical Human Orchestration in Smart Cities,” in IEEE International Conference on Smart Computing (SMARTCOMP), pp. 115-120, 2019.

6. **N. Patrizi**, S.K. LaTouf, E.E. Tsiropoulou, and S. Papavassiliou, ”Prosumer-centric Self-sustained Smart Grid Systems” in IEEE Systems Journal (to appear)

The research work under review are listed below:

1. **N. Patrizi**, S.K. LaTouf, E.E. Tsiropoulou, and S. Papavassiliou, ”Museum and Visitor Interaction & Feedback Orchestration Enabled by Labor Economics,” in IEEE Transactions on Computational Social Systems. (major revision)

1.3.3 Organization

This section summarizes the main structure and organization of the rest of the document.

Chapter 2 presents a socio-physical human orchestration framework in smart cities based on game theory and reinforcement learning. The efficient management of a smart city and the improvement of the quality of humans' every-day life are becoming challenging problems due to smart cities' increased heterogeneity and complexity. In this chapter, we present a novel socio-physical human orchestration framework to deal with the aforementioned issues, by capitalizing on recent advances in game theory and reinforcement learning. Initially, each human selects, in a distributed manner, a Point of Interest (PoI) that it wants to visit, by acting as stochastic learning automaton, exploiting the socio-physical conditions of the environment while learning from its previous experiences. As a result, those humans that have selected a specific PoI to visit, "compete" with each other in order to finally perform their visit. The humans' behavior is studied as a non-cooperative game among them, via adopting the theory of minority games, while the concluding Nash equilibrium point identifies the humans that will finally visit each PoI. A low complexity algorithm is introduced to realize the overall framework, while the performance of the proposed approach is evaluated through modeling and simulation under several scenarios, and its superiority is demonstrated.

Chapter 3 proposes a socio-aware public safety framework design based on a contract-theoretic approach. Given the substantial penetration of social networks in citizens' everyday life activities, the success of a public safety system depends on the citizens' incentivization by the Emergency Control Center (ECC), and their effective effort contribution in the overall disaster management operation. In this chapter, we introduce a formal method based on the principles of Contract Theory, to identify the optimal rewards to the citizens from the ECC's perspective, and the optimal invested effort from the citizens' side, referred to as contract pairs. The identification of these contract pairs (i.e., rewards and respective efforts) between the ECC and each citizen, depend on each citizen's social and communication characteristics that are used to define their specific type and profile, while they are properly reflected in the corresponding designed utility functions to be optimized. The problem under consideration is treated for both cases of complete (ideal) and incomplete (realistic) information availability, with respect to the level of knowledge of the ECC about the exact type of each citizen. The overall framework was evaluated via modeling and simulation, in terms of its efficiency and effectiveness, by studying multiple operation approaches and scenarios.

Chapter 4 introduces a Unmanned Aerial Vehicle or UAV-enabled dynamic multi-target tracking and sensing framework. In this chapter, initially, a holistic reputation model is introduced to evaluate the targets' potential in offloading useful data to the UAVs. Based on this model, and taking into account UAVs and targets tracking and sensing characteristics, a dynamic intelligent matching between the UAVs and the targets is performed. In such a setting, the incentivization of the targets to perform the data offloading is based on an effort-based pricing that the UAVs offer to the targets. The emerging optimization problem towards determining each target's optimal amount of offloaded data and the corresponding effort-based price that the UAV offers to the target, is treated as a Stackelberg game between each target and the associated UAV. The properties of existence, uniqueness and convergence to the Stackelberg Equilibrium are proven. Detailed numerical results are presented highlighting the key operational features and the performance benefits of the proposed framework.

Chapter 5 presents a contract-theoretic resource control in wireless powered communication public safety systems. Recent technological advances in the use of UAVs and Wireless Powered Communications (WPC) have enabled the energy efficient operation of the Public Safety Networks (PSN) during disaster scenarios. In this chapter, an energy efficient information flow and energy harvesting framework capturing users' risk-aware characteristics is introduced based on the principles of Contract Theory. To better support the operational effectiveness of the proposed framework, users are clustered in rescue groups following a socio-physical-aware group formation mechanism, while rescue leaders for each group are selected. A reinforcement learning approach is applied to enable the optimal matching between the UAVs and the rescue leaders in a distributed and efficient manner. The proposed contract-theoretic framework models the UAVs-victims relation based on a labor market setting via offering rewards to the users (incentives) in order to compensate them for their invested labor (reporting information). Detailed numerical results demonstrate the benefits and superiority of the proposed framework under different settings.

Chapter 6 studies the health data acquisition from wearable devices during a pandemic by following a techno-economics approach. In this chapter, we introduce a behavioral and labor economics based approach to address the challenge of citizens' health data acquisition during a pandemic, in a smart city scenario consisting of the healthcare operator, multiple businesses, and citizens with wearable devices. Initially, a reinforcement learning approach is adopted in order for the citizens to select the business to visit, exploiting both social and physical characteristics of all involved entities. Subsequently, following the principles of behavioral economics, the problem of the citizens' incentivization by the businesses to provide their health data via offering personalized rewards is studied. The solution of the corresponding optimization problem concludes to a contract between the business and each citizen associated with this business, containing the optimal reward and optimal portion of reported data. The process is completed by introducing an optimization framework, where the healthcare operator incentivizes the businesses to provide the collected health data to it, by providing them tailored rewards. This is founded on the principles of Contract Theory, where the health-care operator aims at maximizing its benefit from the data acquisition process, while guaranteeing that the optimal determined contracts are acceptable by the respective businesses. Finally, through modeling and simulation, the performance, effectiveness, and robustness of the overall proposed framework is demonstrated, under various realistic scenarios.

Chapter 7 evaluates a contract-theoretic model to enable visitors of museum to provide feedback to the museum in a fairly-incentivized manner. In this chapter, we address the problem of modeling and orchestrating the interactions between a museum and its visitors, viewing the system as a CPSS. In particular, the museum operator provides monetary rewards to the visitors in exchange for their contributions, which are expressed as their total number of provided feedback evaluations of visited exhibits over their touring time. The interactions among the museum operator and visitors are captured in appropriately designed utility functions following the principles of labor economics, while the visitors' behavioral characteristics are utilized to define their unique types. Under such a setting and formulation, the goal of the museum operator is to optimize their profit and benefits, while jointly satisfying the visitors' quality of experience prerequisites, as reflected via their utility functions. The corresponding optimization problem is treated and solved under the general and realistic case of incomplete information, wherein the museum operator estimates the visitors' types probabilistically. The resulting outcome, referred to as "optimal contract" jointly determines the visitors' optimal contributions, as well as the museum operator's optimal amount of personalized rewards provided to each visitor. The performance of the proposed approach is evaluated through modeling and simulation, and detailed numerical results are presented to demonstrate the key benefits of the proposed optimization approach, versus either type-agnostic or heuristic alternatives.

Chapter 8 presents a prosumer-centric self-sustained Smart Grid system. Modern Smart Grid systems exploit a two-way interaction paradigm between the utility and the electricity user and promote the role of prosumer, as a new user type, able to generate and sell energy, or consume energy. Within such a setting, the prosumers and their interactions with the microgrid system become of high significance for its efficient operation. In this chapter, to model the corresponding interactions, we introduce a labor economics-based framework by exploiting the principles of Contract Theory, that jointly achieves the satisfaction of the various interacting system entities, that is the Microgid Operator (MGO) and the prosumers. The MGO offers personalized rewards to the sellers and buyers, to incentivize them to sell and purchase energy, respectively. To provide a stable and efficient operation point, while aiming at jointly satisfying the profit and requirements of the involved competing parties, optimal personalized contracts, i.e., rewards and amount of sold/purchased energy, are determined, by formulating and solving contract-theoretic optimization problems between the MGO and the sellers or byers. The analysis is provided for both cases of complete and incomplete information availability regarding the prosumers' types. Detailed numerical results are presented to demonstrate the operation characteristics of the proposed framework under diverse scenarios.

Lastly, **Chapter 9** concludes the Ph.D. dissertation with an overall summary of the content and a review of its contributions. Additionally, a segment is devoted to the presentation of potential future research directions stemming from this work.

Chapter 2

Human Orchestration in Smart cities

2.1 Introduction

Recent years have witnessed the rapid growth of smart cities which, among other benefits, provide smart service systems to enrich and support people's lives and entertainment options. [54]. People can join different social events (e.g., dining out, playing sports) by visiting different Points of Interest (PoIs), e.g., restaurants, stadiums, tax offices, in their daily life and decide which places to go to according to some social and physical parameters (e.g., location preferences, geographical proximity). The efficient orchestration of humans within a smart city can result in many fold benefits and catalyze the sustained economic growth of the smart city. However, the tremendous increase in available information for decision-making, the large number of possible PoIs within a smart city along with specific social and physical characteristics and constraints, makes the problem of selecting the most interesting PoI and deciding whether to visit it, extremely challenging.

2.1.1 Related Work & Motivation

Recently, a number of research works have been proposed in the context of planning PoI visits, mainly exploiting the information extracted from the event-based social networks (EBSN) such as Foursquare, Meetup, and Twitter. In [55], the authors analyze the humans' behavior in EBSNs by exploiting their social activities and interactions towards explaining their attendance in PoIs and identifying the most influential factors on the humans' decisions. This study has been extended in [56], where the authors provide a similar analysis, regarding groups of humans who belong to common social groups, by utilizing a Mixed Markov Model to identify the groups' behavioral patterns. In [57], the authors introduce various recommendation algorithms of PoIs to be visited by the humans based on their past visited PoIs, the physical location of the available PoIs, the social interaction among the humans and their similarity among each other. A traveling recommender system is proposed in [58], by jointly considering the PoIs popularity, the similarity of the humans that visit the same PoI, and the similarity of the available PoIs towards recommending PoIs.

Furthermore, in [59], the authors study the problem of real-time PoI and event recommendations to the humans by introducing the event-participant arrangement strategy. Following this concept, the humans' satisfaction scores, regarding an arrangement of visiting a PoI, are updated in real-time and the humans can accept or reject the proposed arrangement. A human-centric approach is also followed in [60], where the humans are assigned to PoIs and events aiming at maximizing the humans' perceived satisfaction. On the other hand, a system-centric approach is proposed in [61] to support the PoIs' management towards maximizing their perceived "satisfaction", which is expressed in terms of revenue and publicity. A more holistic approach is introduced in [62] by exploiting the whole set of EBSNs functionalities to recommend PoIs to humans, social groups to humans, and tags to groups.

As it becomes apparent from the above discussion, several studies have constructed models of recommending PoIs to the humans, either by following a human-centric or a system-centric approach. Furthermore, the literature is already mature enough in exploiting the information available in EBSNs, such as humans' interests in PoIs, humans' social interactions, geographical proximity to the PoIs, etc. However, to the best of our knowledge, no prior work has dealt with the problem of socio-physical autonomous human orchestration in a smart city environment, where humans can exploit their personal social and physical characteristics, as well as those of the PoIs to make efficient distributed and autonomous decisions that improve their personal reward from the visited PoIs.

2.1.2 Contributions & Outline

Our research work aims exactly at filling the aforementioned research gap and proposes a holistic human-centric distributed approach realizing (i) the PoI selection by the humans, via a reinforcement learning technique, and (ii) the human's decision-making process of visiting a PoI, by exploiting the theory of minority games. Our proposed framework consists of two layers to treat the socio-physical autonomous human orchestration in a smart city. At the first layer, the humans are considered as stochastic learning automata who learn from their past choices of PoIs and the reaction of the smart city environment towards selecting a PoI that will improve their experienced reward. The humans make probabilistic choices of PoIs until they reach a firm PoI selection by exploiting their social characteristics, e.g., interest to visit a PoI, social interaction among the humans that visit the same PoI, and the physical characteristics, e.g., cost of visit, physical proximity to the PoI, experienced Quality of Service (QoS) from visiting the PoI, PoI's capacity and availability.

Given the convergence of the humans' PoI selection, the humans that selected the same PoI and expressed their initial interest to visit it, "compete" with each other towards finally visiting the PoI and improving their experienced reward from their visit. The latter humans' behavior and interaction is modeled as a non-cooperative game among the humans that selected the same PoI towards determining their final attendance or not. Towards showing the existence of the game's Nash equilibrium, which identifies the specific humans who will visit the PoI, the theory of minority games is adopted. A distributed and low-complexity algorithm is introduced, which determines both the humans' PoI selection and the humans who visit the PoIs. Detailed numerical and comparative results demonstrate that the proposed holistic framework concludes to a promising solution for realizing the autonomous human orchestration in a smart city, that conforms with the needs and requirements of both the humans and the smart city planning and management.

The rest of the chapter is organized as follows. In Section 2.2, the overall system model is described, while in Section 2.3, our proposed human-centric reinforcement learning-based PoI selection process is presented. Section 2.4 introduces the autonomous human orchestration to the PoIs based on the theory of minority games, while in Section 2.5 the Smart Orchestration in Points of Interest (SmartPoI) algorithm is presented. Finally, a detailed numerical evaluation of our approach via modeling and simulation is presented in Section 2.6, while Section 2.7 concludes the chapter.

2.2 System Model

In this chapter, a smart city environment is considered, with humans interested in visiting various PoIs inside of a smart city. These humans will decide which PoI they are interested in visiting by taking into account the socio-physical characteristics of each PoI and the overall system. Once they have a choice on which PoI they are

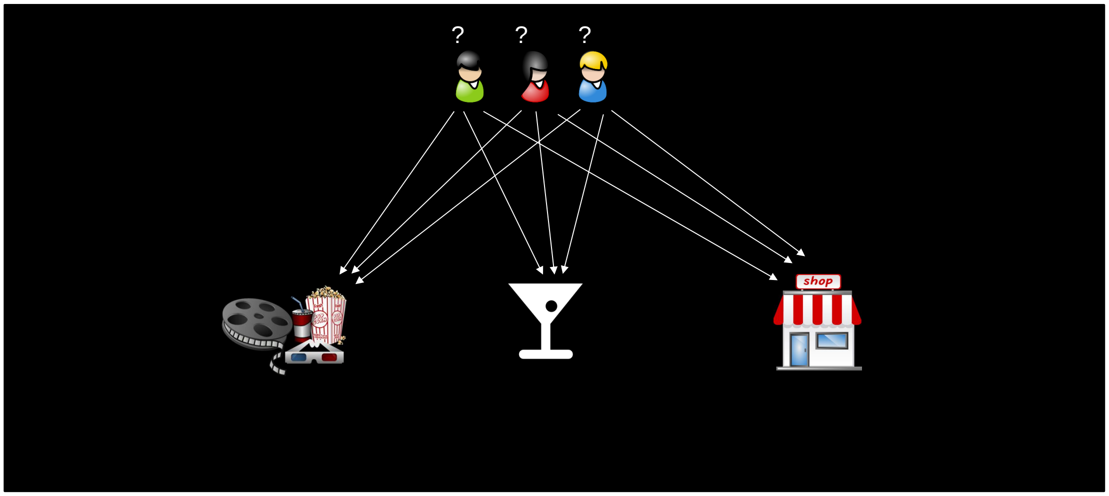

Figure 2.1: Humans autonomous decision making regrading the Point of Interest that they will visit.

interested in visiting they each play a minority game to determine which humans will actually visit the PoI. The humans of the system are dentoted as $|N|$, with the humans residing in the smart cities boundaries, with the set denoted as $N = \{1, ..., n, ..., |N|\}$. These humans select from the various PoIs $|S|$ (e.g., restaurants, theaters, tax offices, police station, etc), with the corresponding set of PoIs being denoted as $S = \{1, ..., s, ..., |S|\}$. Each human will make selection to as to which PoI they have a desire to visit based on their own personal social characteristics, as well as the physical conditions that are available in the smart city environment, as shown in Figure 2.1.

The physical and social parameters considered in the system are designed to provide a holistic view of what aspects matter and the importance thereof to each human, with each human having personalized preferences. The first parameter considered is the interest of the human to visiting a specific PoI. For example, a human might need to pickup groceries in order to make dinner and would thus have a high interest in visiting a grocery store. Specifically, each human $n, n \in N$ has a personal interest $i_{n,s}$ to visit a PoI $s, s \in S$. The interest degree $i_{n,s}$ ranges from zero to one, i.e, $i_{n,s} \in [0,1]$, with smaller values representing less interest to visit the PoI and larger values meaning a greater interest to visit that PoI. This shows and considers the unique preference of a human to a space based on their own personal view of the space. Next, as many places that humans visit are based off of a social element (e.g., restaurant, movie theater, etc.), the social aspect of a PoI should be considered. Naturally, this social aspect is based upon personal preferences of one human's view towards other humans in the space. For instance, a human's social interest for visiting a PoI with humans that they enjoy interacting with will have a higher social interest value compared to that of a PoI with humans they don't like interacting with or a PoI with no humans in it. Thus, the peers' influence on visiting a PoI is captured by the social interest $SI_{n,j}$ which expresses the level of willingness of humans n, j to socially interact with each other. We set the range of $SI_{n,j}$ as $SI_{n,j} \in [0,1]$ and we assume that the level of social interaction among two humans $n, j \in N$ is directly proportional to the value of $SI_{n,j}$.

The humans and the PoIs in a smart city are characterized by some physical conditions and parameters. Each PoI s has an associated cost c_s to serve the needs of the humans. For example, there are restaurants that are more expensive compared to others which has a drastic influence on a humans' decisions to visit them. The PoI's cost c_s of serving a human is normalized with respect to the maximum cost of a PoI in the smart city, i.e., $c_s \in [0,1]$, with values of c_s closer to 1 being more expensive. Naturally, each PoI has a limited amount of humans that can be accommodated at any time, thus every PoI is characterized by a physical capacity N_s^{thres}, $s \in S$ of humans that it represents the amount of people that can be served. For example, restaurants will have a limited number of seats and other PoIs will have a limited based off of fire safety laws as well. Furthermore, humans will tend to not want to travel long distances to visit a PoI, consequently the distance $d_{n,s}$ of human n from the PoI s also plays a role in the human's personal physical factor that weighs on their corresponding decision with regard to which PoI they would like to visit. In our analysis, the distance $d_{n,s}$ is also normalized with respect to the maximum, thus, $d_{n,s} \in [0, 1]$ with values close to zero meaning that the human is close to that respective PoI. Additionally, humans will prefer PoIs where they are efficiently and effectively served. PoIs with more people in them are congested and will cause humans to have a probability of not being served effectively. This shows that the number of humans $|N|_s^{Go}$, who decide to visit a PoI s is a critical factor in the humans' decision to go to a PoI. This leads to the defined experienced Quality of Service (QoS) of human n by visiting a PoI s which is denoted by $QoS_{n,s}$, with $QoS_{n,s} \in [0, 1]$. The overall Quality of Service is directly proportional to the value of $\sum_{k=0}^{t} QoS_{n,s}^{[k]}$, which expresses the human's cumulative experienced QoS over the time including all the PoIs that the human has visited. If the humans that go to a PoI are more than the PoI's capacity, then their experienced QoS is zero (as no human would be served effectively), i.e., $QoS_{n,s} = 0$, if $|N|_s^{Go} > N_s^{thres}$, while if the number of humans that visit the PoI is less than the capacity of the PoI, the normalized human's QoS is given by $QoS_{n,s} = 1 - \frac{|N|_s^{Go}}{N_s^{thres}}$, if $|N|_s^{Go} \leq N_s^{thres}$.

2.3 Socio-physical Point of Interest selection

In this section, our goal is to devise a distributed and autonomous mechanism to enable the humans to select which PoIs they are potentially interested in visiting based on the socio-physical characteristics previously mentioned. To accomplish this, we utilize a reinforcement learning technique, which allows for the humans to learn from their prior choices and the effect that it produced on the overall smart city. The humans are considered as stochastic learning automata [18] and at each time slot t of the reinforcement learning loop, they select to visit a PoI from their available set of actions $a_n(t) = \{a_1, ..., a_s, ..., a_{|S|}\}$, which represents the available PoIs within the smart city. The physical meaning of the time slot t can be defined based on the specific smart city application. Towards selecting a PoI, the humans consider their social and physical characteristics (Section 2.2): (i) N_s^{thres}: the maximum number of humans that the PoI s can accommodate, (ii) c_s: the normalized cost associated with visiting and being served at the PoI s, (iii) $d_{n,s}$: the normalized physical distance of human n from PoI s, (iv) $i_{n,s}^{[t]}$: the normalized interest of human n to visit the PoI s at time slot t, (v) $|N|_s^{Go[t]}$: the number of humans that have selected to go to the PoI s at the time slot t, (vi) $\sum_{j=1}^{|N|_s^{Go[t]}} SI_{n,j}^{[t]}$: the total social interest and interaction of human n with all the other humans $|N|_s^{Go[t]}$ that have selected to go to the PoI s at time slot t, and (vii) $\sum_{k=0}^{t} QoS_{n,s}^{[k]}$: the cumulative QoS that the human n has experienced until the time slot t including all the PoIs that the human has visited.

By combining the above humans' social characteristics and PoIs' physical parameters, we define the reward function that a human n experiences by visiting a PoI s, as follows.

$$r_{n,s}^{[t+1]} = \frac{N_s^{thres} \cdot i_{n,s}^{[t]} \cdot \sum_{j=1}^{|N|_s^{Go[t]}} SI_{n,j}^{[t]} \cdot \sum_{k=0}^{t} QoS_{n,s}^{[k]}}{|N|_s^{Go[t]} \cdot c_s \cdot d_{n,s}} \tag{2.1}$$

The reward function $r_{n,s}^{[t+1]}$ is dynamically determined by the human's past experience (e.g., $\sum_{k=0}^{t} QoS_{n,s}^{[k]}$), as well as by the reaction of the smart city environment, meaning the choices of the rest of the humans residing in the smart city. Also, the reward function $r_{n,s}^{[t]}$ of each human n per available PoI s is normalized as $\tilde{r}_{n,s}^{[t+1]} = \frac{r_{n,s}^{[t+1]}}{\sum_{s \in S} r_{n,s}^{[t+1]}}$ to represent the reward probability $\tilde{r}_{n,s}^{[t+1]}$, $0 \leq \tilde{r}_{n,s}^{[t+1]} \leq 1$ of the human n per each PoI s. In a nutshell, the reward probability $\tilde{r}_{n,s}^{[t+1]}$ reflects the potential satisfaction that the human n may experience by visiting the PoI s at time slot t. A graphical representation of the reward function and its individual components is presented in Figure 2.2. The humans consider their reward probabilities in order to determine

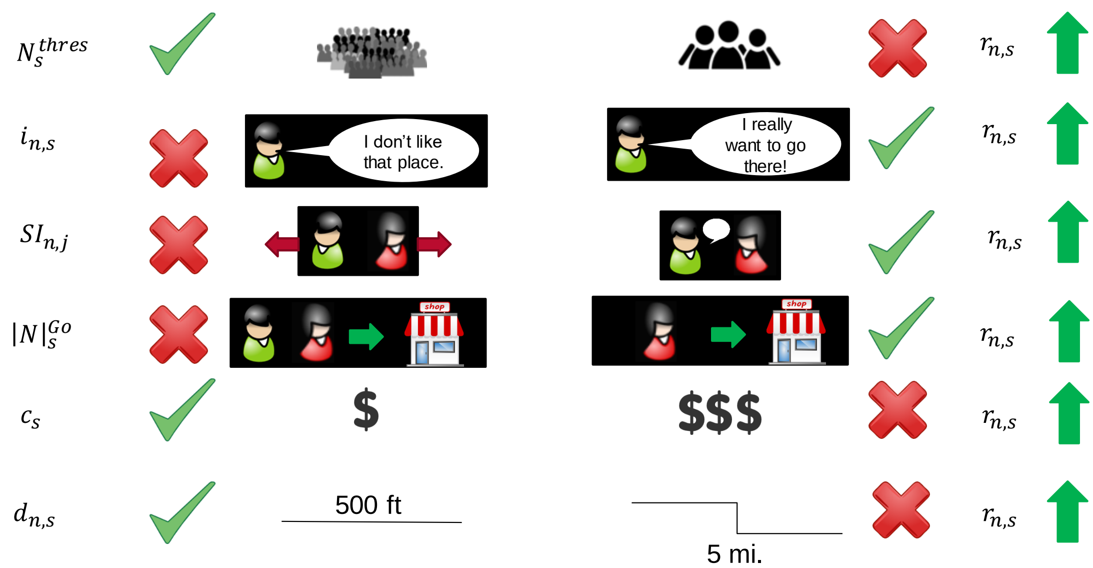

Figure 2.2: Graphical representation of the reward function.

and update their action probabilities of selecting a PoI.

Each human acts as a stochastic learning automaton and updates its action probability vector $\mathbf{Pr_n^{[t]}} = [Pr_{n,1}^{[t]}, ..., Pr_{n,s}^{[t]}, ..., Pr_{n,S}^{[t]}]$, where $Pr_{n,s}^{[t]}$ represents the probability that the human n will select the PoI s at time slot t. Based on the theory of the stochastic learning automata [18, 63, 64], the humans update their action

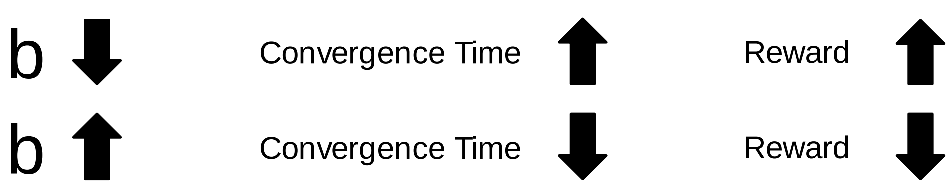

Figure 2.3: Graphical representation of the dependence of the convergence time and the corresponding achieved reward from the learning rate b.

probabilities based on the following rule [65].

$$Pr_{n,s}^{[t+1]} = Pr_{n,s}^{[t]} + b \cdot \tilde{r}_{n,s}^{[t]} \cdot (1 - Pr_{n,s}^{[t]}), s_n^{[t]} = s_n^{[t+1]} \tag{2.2a}$$

$$Pr_{n,s}^{[t+1]} = Pr_{n,s}^{[t]} - b \cdot \tilde{r}_{n,s}^{[t]} \cdot Pr_{n,s}^{[t]}, s_n^{[t]} \neq s_n^{[t+1]} \tag{2.2b}$$

where $0 \leq b \leq 1$ represents the humans' learning rate in terms of exploiting the smart city environment. The dependence of the convergence time and the corresponding achieved reward from the learning rate b is presented in Figure 2.3. The human's probability to select the same PoI in the next time slot $t+1$ is updated following Eq. 2.2a, while the human's probability to select a different PoI in the next time slot $t+1$ is calculated by Eq. 2.2b. Also, it is noted that the humans have initially no prior knowledge regarding their action probabilities, thus the initial selection of a PoI by the humans is made with equal probability, i.e., $Pr_{n,s}^{[t=0]} = \frac{1}{|S|}, \forall s \in S$. The algorithmic description of the socio-physical PoI selection based on the proposed reinforcement learning technique and the convergence of the humans' action probabilities are studied in Section 2.5.

2.4 Autonomous Human Orchestration based on Minority Games

After the socio-physical PoI selection by the humans, a number of humans $|N|_s$ has selected to potentially visit the PoI s at the next time slot, where $N_s = \{1, ..., |N|_s\}$ denotes their corresponding set. The humans "compete" with each other towards finally visiting the PoI that they have initially selected. The interactions and behavior of the humans, who through the reinforcement learning framework expressed interest in visiting the same PoI, is further captured via a non-cooperative game among them. Specifically, the theory of minority games is adopted, which proposes that a number of players (i.e., humans) repeatedly compete with each other to be in the minority group via making an action of the two available ones, i.e., go or not to the initially selected PoI. At each iteration *ite* of the game, the humans that belong to the minority group perceive increased satisfaction and they promote their winning strategy for the next iteration of the game. The main benefit of the minority games is that they have a non-empty set of Pure Nash equilibria (PNE) [18].

Let us denote the minority game as $G_{MG} = [N_s, \{A_n\}, \{f_{a_n}(n)\}]$, where N_s is the set of humans that have selected to visit the PoI s following the reinforcement learning framework (Section 2.3). At each iteration *ite* of the minority game, each human can decide to visit the PoI ($a_n^{ite} = 1$) or not ($a_n^{ite} = 0$). The set of human's strategies is denoted as $A_n = \{0, 1\}$, $a_n^{ite} \in A_n$. For each strategy $a_n^{ite} \in A_n$, there is a payoff function $f_{a_n}^{ite} : \{1, ..., n, ..., |N|\} \to \mathbb{R}$, which represents the reward that the human n experiences by making the action a_n at the iteration *ite* of the minority game. The payoff function $f_{a_n}^{ite}$ is formulated as follows.

$$f_{a_n}^{ite} = \begin{cases} 1, \text{if } |N|_s^{a_n} \leq N_s^{thres} \\ 0, \text{otherwise} \end{cases} \tag{2.3}$$

where if the number of humans that select a strategy a_n (i.e., $|N|_s^{a_n}$) is less than the physical capacity N_s^{thres} of the PoI s then they promote their action, i.e., $f_{a_n}^{ite} = 1$. To solve the minority game and determine its Pure Nash equilibrium, a distributed learning algorithm is required. This goal can be achieved by multiple distributed learning techniques, e.g., Q-learning, exponential learning, trial and error learning. In this chapter, we have adopted an exponential learning technique to determine in an autonomous and distributed manner the Pure Nash equilibrium of the minority game G_{MG} (see Section 2.5).

2.5 Smart Orchestration in Points of Interest (SmartPoI) Algorithm

In this section, the distributed Smart Orchestration in PoIs (SmartPoI) algorithm is presented. At each time slot t, each human n acts as a stochastic learning automaton making its choice of the PoI that wants to visit, based on its action probabilities $\mathbf{Pr_n^{[t]}} = [Pr_{n,1}^{[t]}, \dots, Pr_{n,|S|}^{[t]}]$. After each human's choice, a cluster of humans $|N|_s, \forall s \in S$ is constructed, and a minority game is played to determine the set of humans who finally visit the PoI (N_s^{GO}), and the corresponding set of humans who do not visit the PoI (N_s^{NGO}). For the minority game played for each PoI s, a distributed exponential learning algorithm is adopted, which leads the humans to make smart choices by considering only their past actions and converge to one of the $\binom{N_s}{N_s^{thres}-1} + \binom{N_s}{N_s^{thres}+1}$ PNE points [66]. For each cluster of humans $|N|_s$ that selected the PoI s, each human n by starting with equal probabilities of going and not going, i.e., $pr^0_{n,a_n=0} = pr^0_{n,a_n=1} = 0.5$, and zero scores, i.e., $\pi^0_{n,a_n=0} = \pi^0_{n,a_n=0} = 0$, at each iteration ite of the minority game the human n determines its action a_n^{ite} and regarding its payoff $f_{a_n}^{ite}$ (Eq. 2.3) and the winning action w^{ite}, it updates its chosen action's score π^{ite}_{n,a_n}. It is highlighted that the winning action w^{ite} is evaluated regarding the winning minority group. Then, each human n evaluates its next time slots' reward probability $\tilde{r}_{n,s}^{[t+1]}, \forall s \in S$, and updates its action probabilities $Pr_{n,s}^{[t+1]}, \forall s \in S$ (Eq. 2.2a, 2.2b).

Regarding the SmartPoI algorithm's complexity, at each time slot t of the stochastic learning automata, the minority games at all PoIs are played in parallel. Moreover, since the complexity of each minority game is $\mathcal{O}(|N|_s)$, by denoting as Ite the number of iterations that are needed for the convergence of the minority game that finishes last, the overall complexity of all the minority games is $\mathcal{O}(Ite \cdot |N|)$. Furthermore, since the evaluation of the reward probability and the update of the action probabilities of each human n for each PoI s, is performed in a constant time,

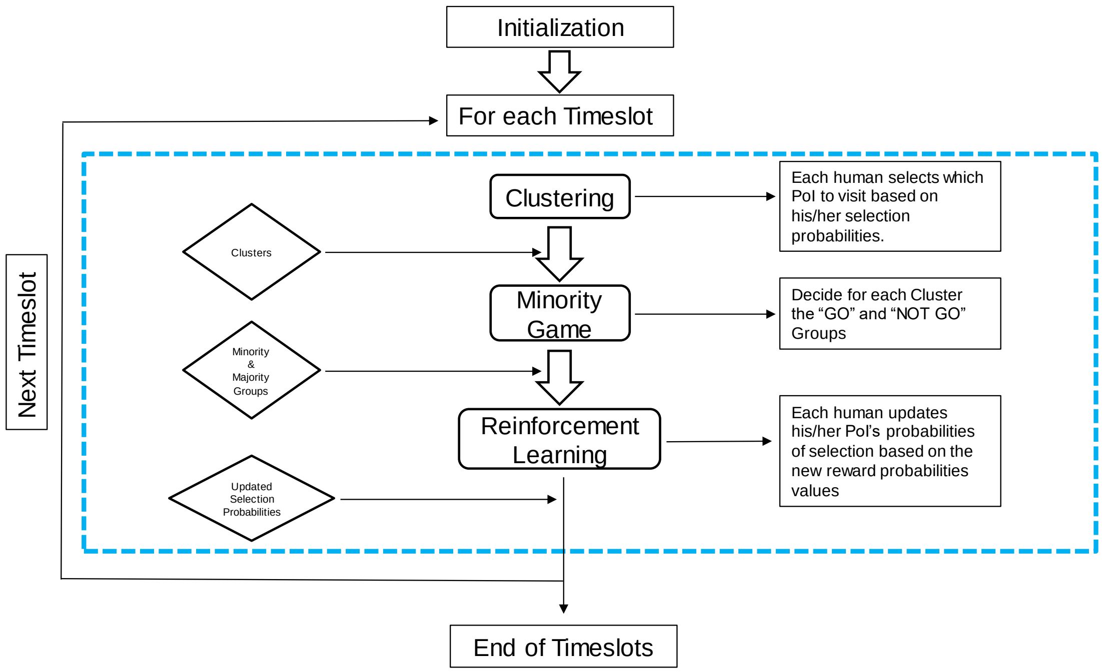

Figure 2.4: Graphical representation of the overall proposed framework in this research work summarizing the flow of information, as well as of the control actions to conclude to the autonomous decision making process.

the complexity of the rest part of the SmartPoI algorithm is $\mathcal{O}(|N| \cdot |S|)$. Finally, by denoting as T the numbers of time slots that are needed for the convergence of the stochastic learning automata, the overal complexity of the SmartPoI algorithm is $\mathcal{O}(T \cdot (Ite \cdot |N| + |N| \cdot |S|))$.

A graphical representation of the overall proposed framework in this research work is presented in Figure 2.4, summarizing the flow of information, as well as of the control actions to conclude to the autonomous decision making process.

Algorithm 1 SmartPoI Algorithm

Function {Main}:

Input/Initialization: $N, S, i_{n,s}^{[0]}, d_{n,s}, SI_{n,j}, c_s, N_s^{thres}$
$t = 0, Conv = 0, Pr_{n,s}^{[0]} = \frac{1}{|S|}, \forall n \in N, \forall s \in S$

Output: $\mathbf{Pr}^* = [\mathbf{Pr_1^*}, \ldots, \mathbf{Pr_n^*}, \ldots, \mathbf{Pr_{|N|}^*}]$

while $Conv == 0$ **do**

$\quad N_s = \varnothing, \forall s \in S$

$\quad$ *Choose* $a_n(t) \in S, N_{a_n(t)} = N_{a_n(t)} \cup \{n\}, \forall n \in N$
$\quad$ *based on* $\mathbf{Pr_n^{[t]}}$

$\quad MinorityGame(N_s, N_s^{thres}), \forall s \in S$

$\quad$ *Evaluate* $r_{n,s}^{[t+1]}, \tilde{r}_{n,s}^{[t+1]}, Pr_{n,a_n(t)}^{[t+1]}, Pr_{n,s}^{[t+1]}$ *via Eq. 2.1, 2.2a, 2.2b* $\forall n \in N, \forall s \in S$

$\quad Conv = 1$, ***if*** $\forall n \in N$, $\exists s \in S$: $|Pr_{n,s}^{[t+1]} - 1| \leq 0.99$

$\quad t = t + 1$

end while

$\mathbf{Pr_n^*} = \mathbf{Pr_n^{[t]}}, \forall n \in N$

EndFunction

Function {MinorityGame}:

Input/Initialization: $N_s, N_s^{thres}, pr_{n,a_n}^{ite} = 0.5, \pi_{n,a_n}^{ite} = 0, ite = 0, Conv = 0, \forall n \in N, \forall a_n \in A_n$

Output: N_s^{GO}, N_s^{NGO}

while $Conv == 0$ **do**

$\quad N_s^{GO} = N_s^{NGO} = \varnothing$

$\quad$ *Choose* $a_n^{ite}, \forall n \in N_s$ *based on* $pr_n^{ite} = [pr_{n,0}^{ite}, pr_{n,1}^{ite}]$

$\quad$ ***if*** $a_n^{ite} = 1$, ***then*** $N_s^{GO} = N_s^{GO} \cup \{n\}$
$\quad$ ***else*** $N_s^{NGO} = N_s^{NGO} \cup \{n\}$

$\quad$ ***if*** $|N_s^{GO}| \leq N_s^{thres}$, ***then*** $w^{ite} = 1$ ***else*** $w^{ite} = 0$

$\quad \pi_{n,a_n}^{ite+1} = \pi_{n,a_n}^{ite} + f_{a_n}^{ite}$

$\quad pr_{n,a_n}^{ite+1} = \exp(\gamma \cdot \pi_{n,a_n}^{ite+1}) / \sum_{\forall a_n \in A_n} \exp(\gamma \cdot \pi_{n,a_n}^{ite+1})$ $\forall a_n \in A_n$, $\forall n \in N_s$

$\quad Conv = 1$, ***if*** $\forall n \in N_s$, $\exists a_n \in A_n$: $|pr_{n,a_n}^{ite+1} - 1| \leq 0.99$

$\quad ite = ite + 1$

end while

EndFunction

2.6 Experiments

2.6.1 Experiment Setup

In this section, a detailed numerical evaluation of the proposed approach is presented in terms of the overall framework's operation efficiency (Section 2.6.2) and superiority compared to other alternatives (Section 2.6.3). For our simulations, that were carried out using MATLAB software, we considered a smart city area that consists of $|N| = 100$ humans randomly distributed in the smart city setting and $|S| = 6$ PoIs. The interest $i_{n,s}$ as well as the social interest of interaction among the humans $SI_{n,j}$ are randomly and uniformly assigned to the humans, while $\mathbf{N^{thres}} = [6, 8, 10, 12, 14, 16]$ and $\mathbf{c} = [0.166, 0.333, 0.5, 0.666, 0.833, .999]$. A detailed Monte Carlo analysis has been executed for all the presented numerical results considering averages over $10,000$ executions.

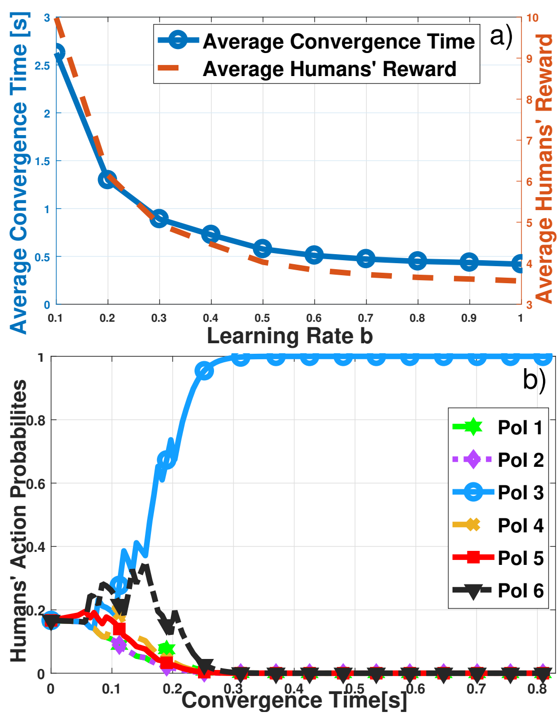

Figure 2.5: (a) Average convergence time and average humans' reward vs b and (b) Action probabilities convergence

2.6.2 SmartPoI Framework's Operation

First, we evaluate the operation of the socio-physical PoI selection following the proposed reinforcement learning technique. Figure 2.5a presents the impact of the learning rate parameter b on the average convergence time of the PoI selection and the corresponding average reward (Eq. 2.1). The results reveal that for small values of the learning rate parameter, the humans exploit more thoroughly the available PoIs, thus, they make a better choice of PoI, resulting in increased average reward. However, the latter comes with the cost of increased convergence time to a PoI's selection. In the rest of our analysis, we consider $b = 0.4$. Additionally, in Figure 2.5b, the action probabilities convergence is presented for one representative human in the smart city. The results illustrate that the execution time of the proposed PoI selection mechanism is less than 1 sec, which makes it practical for real-life applications.

In Figures 2.6a-2.6d, we present a detailed analysis of the internal operation of the PoI selection reinforcement learning mechanism based on the proposed reward function (Eq. 2.1), which captures humans' and PoIs' social and physical characteristics. Figure 2.6 illustrates the average cluster size of the humans that selected each PoI based on: (a) the varying cost c_s of the PoIs, (b) the varying distance $d_{n,s}, n \in N, s \in S$, (c) the varying PoIs' capacity $N_s^{thres}, s \in S$, (d) all the varying factors of the reward function in Eq. 2.1. It is noted that in Figures 2.6(a)-2.6(c) only one parameter is varying, while the rest of the factors are the same for all the users for all the PoIs for fairness in the comparison. The results reveal that the humans proportionally select the PoI with the lower cost c_s (Figure 2.6a) and the higher capacity (Figure 2.6c). The results also illustrate that the humans select the PoI with the closest physical proximity (Figure 2.6b). In Figure 2.6d a more complex case is examined and presented, where multiple social and physical factors are varying. It is observed that the cost c_s becomes a dominant factor in humans' PoI selection, i.e., more humans select PoIs 1 and 2 which have the relatively lower cost. However, the dominance of the PoIs' cost in the PoI selection can be limited by other factors such as the humans' distance from the PoIs and the PoIs capacity. For example, even if PoI 3 has lower cost compared to PoI 6, less humans select PoI 3, as it has a smaller capacity than PoI 6, thus it can easily become congested and unable to efficiently serve them.

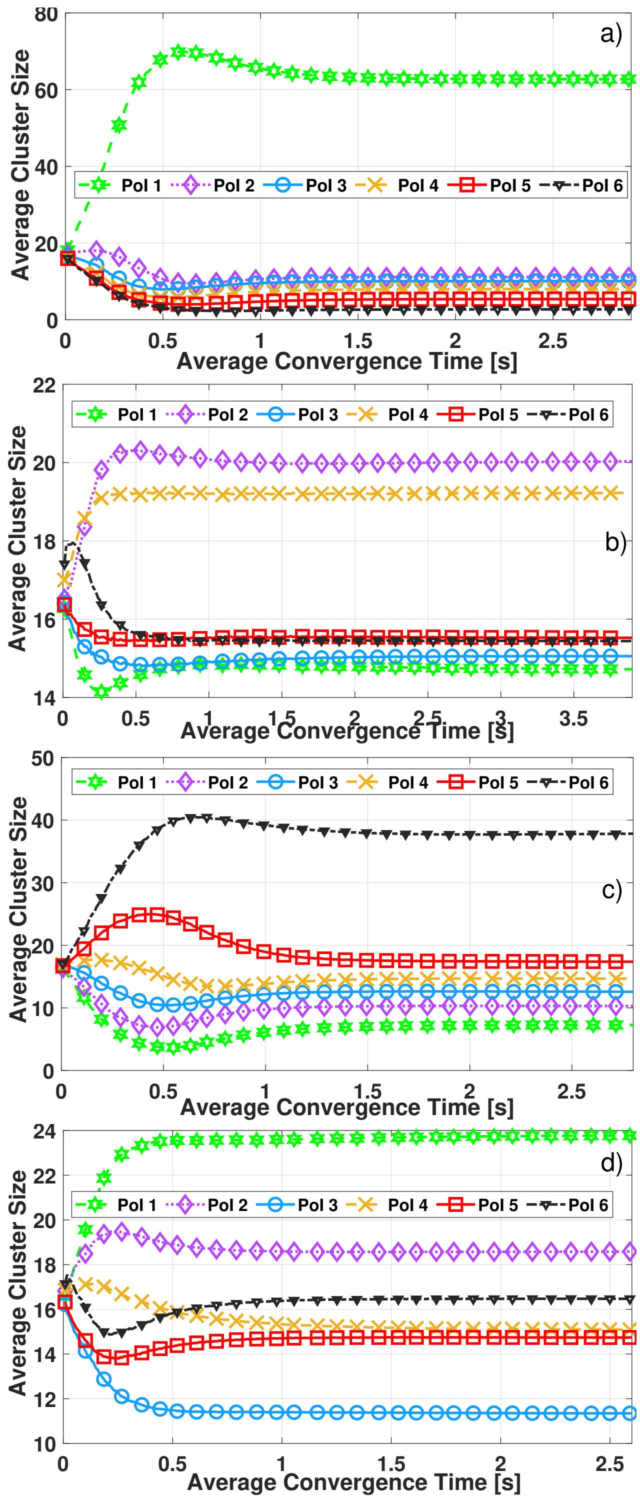

Figure 2.6: Humans' cluster size per PoI for varying (a) PoIs' cost c_s (b) humans' distance from PoIs' $d_{n,s}$, (c) PoIs' capacity N_s^{thres}, and (d) all the socio-physical factors in Eq.2.1

Therefore, we conclude that the holistic consideration of the humans' and PoIs' social and physical characteristics in the PoI selection process can better capture the realistic environment of the smart city.

Next, we discuss the operation of the minority games approach which enables the humans who initially selected a PoI to finally determine if they will visit it. The convergence of the humans' action probabilities is presented in Figure 2.7a for two indicative subjects. Also, Figure 2.7b presents the humans' attendance to one PoI, which has a corresponding capacity $N_{s=2}^{thres} = 8$. The results reveal that the proposed decision-making approach of the minority games is of low time complexity (i.e., order of *msec*) and the number of humans who go to a PoI, stays close to PoI's capacity, thus the PoI serves the humans in an efficient manner.

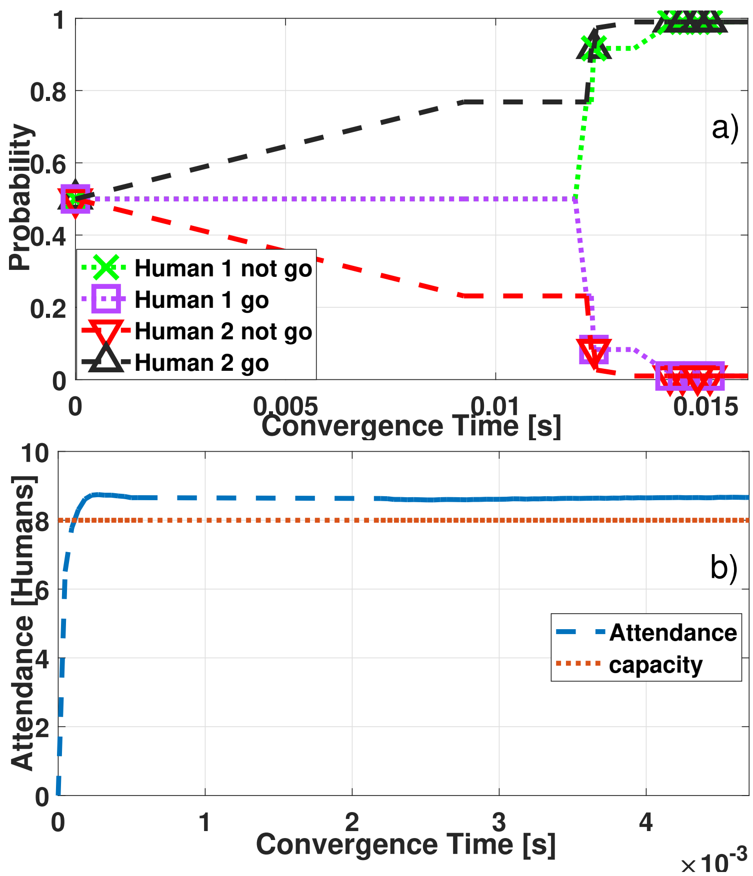

Figure 2.7: Convergence of human's (a) action probabilities and (b) attendance.

2.6.3 Comparative Results

In this subsection, we provide a comparative analysis of our approach focusing on the benefits of: (a) the holistic consideration of the humans' and PoIs' socio-physical characteristics and (b) the stochastic learning automata technique to enable the humans to learn the most beneficial selection of a PoI.

Initially, we consider a scenario, where the PoI selection by the humans and the decision to go to a PoI is performed following the procedure presented in the SmartPoI algorithm, while six different alternatives are examined regarding the considered reward function (Eq. 2.1). In particular, the different cases considered are as follows, (a) cost: $r_{n,s}^{[t+1]} = \frac{1}{c_s}$, (b) distance: $r_{n,s}^{[t+1]} = \frac{1}{d_{n,s}}$, (c) interest: $r_{n,s}^{[t+1]} = i_{n,s}^{[t+1]}$, (d) SmartPoI, i.e., the reward function is given by Eq. 2.1, (e) QoS: $r_{n,s}^{[t+1]} = \sum_{k=0}^{t} QoS_{n,s}^{[k]}$ (f) social interest: $r_{n,s}^{[t+1]} = \sum_{j=1}^{|N|_s^{Go[t]}} SI_{n,j}^{[t]}$. For fairness in the comparison, we use the reward function of Eq. 2.1 to capture the humans' satisfaction (Figure 2.8a). Also the average convergence time to the PoIs selection (Figure 2.8b) and the average cluster size of humans per PoI (Figure 2.8c) are presented. The results reveal that the holistic consideration of the humans' and PoIs' characteristics, i.e., SmartPoI scenario, conclude to improved humans' satisfaction (Figure 2.8a), while allowing the humans to quickly learn their desired PoI selection (Figure 2.8b) and not overcongest the PoIs (Figure 2.8c).

Moreover, the linear relationship of the influential factor (i.e., interest case) with the humans' reward function concludes to a slow update rule of PoI selection and a corresponding low achieved satisfaction compared to a convex relationship (i.e., distance case), which enables the humans to rapidly exploit the smart city environment and make a better PoI selection. Moreover, if the PoI selection is based only on the PoIs' physical characteristics (e.g., cost case), the humans initially select the PoI with the lowest cost (thus, they increase their perceived satisfaction), and when they exceed the PoIs' capacity, they quickly learn that this PoI selection is not beneficial anymore and they choose another PoI.

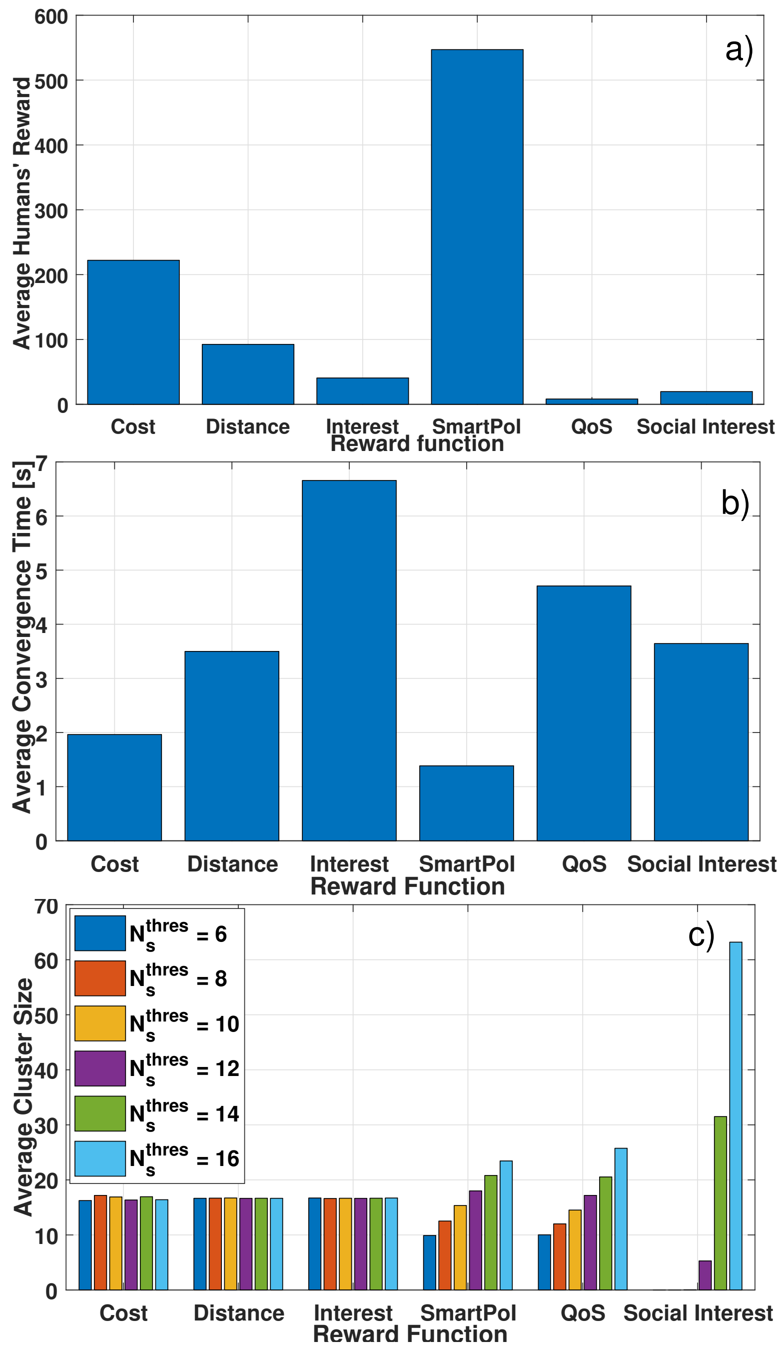

Figure 2.8: Average (a) Humans reward (b) Convergence time, and (c) Cluster size per PoI, for different reward functions

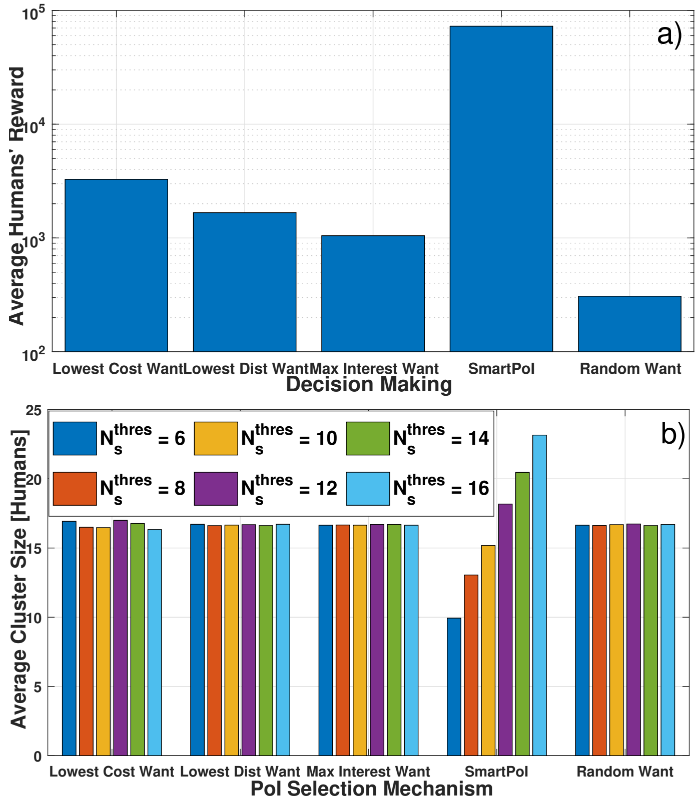

Figure 2.9: Average (a) Humans reward and (b) cluster size per PoI, for different PoIs selection mechanism

Furthermore, if the humans' personal characteristics are considered for the PoI selection, i.e., QoS and social interest cases, the humans have a myopic view of the smart city environment based only on their own perspective, thus they are not able to efficiently and quickly exploit their choices and they achieve low levels of satisfaction. Additionally, in Figure 2.8c it is observed that based on the performed Monte Carlo analysis, the cost, distance, and interest cases conclude to equal human distributions per PoI, while the SmartPoI and QoS cases that consider the PoIs' capacity during the PoI selection process do not overcongest the PoIs. Also, in the social interest case, we observe that the humans tend to select the PoI with the highest capacity, as in this case they have better chances to meet other humans with similar interests.

Next, we consider another comparative scenario, where the humans select to visit a PoI based on the following alternatives: (a) lowest cost, (b) lowest distance from a PoI, (c) maximum interest for a PoI, and (d) randomly, instead of fully exploiting the proposed SmartPoI framework. The results reveal that the PoI selection based on the SmartPoI framework concludes to superior reward for the humans (Figure 2.9a), as they thoroughly exploit their available choices. The random PoI selection gives the worst rewards to the humans, while it is observed that the humans become more satisfied if they pay less to visit a PoI compared to the cases where they have to travel a large distance for their visit or if they are highly interested in visiting the PoI. Finally, following the performed Monte Carlo analysis, the results reveal that the SmartPoI framework does not congest the PoIs, while all the other examined comparative cases equally distribute the humans among the PoIs, thus, congesting some PoIs with small capacity N_s^{thres}.

2.7 Conclusions

In this chapter, the problem of the socio-physical human orchestration in smart city environments is studied by exploiting reinforcement learning and game-theoretic techniques. Initially, the humans act as stochastic learning automata probabilistically selecting to visit a Point of Interest based on the reward that they receive and their past experience. The introduced humans' reward captures their social characteristics, as well as the PoIs' physical characteristics. At the second layer of the proposed approach, the humans that have selected the same PoI "compete" with each other towards finally visiting it. The latter humans' behavior is studied as a non-cooperative minority game among the humans. The Nash equilibrium point of the game is determined, which identifies the specific humans that will finally visit each PoI. A distributed low-complexity algorithm is presented to realize the proposed framework, while the efficiency and superiority of the proposed framework is evaluated and demonstrated through modeling and simulation. Part of our current and future work includes the testing of the proposed framework in the real smart city environment of the City of Albuquerque, New Mexico, USA and based on the realistic outcomes, and observations to fine tune the theoretical model.

Chapter 3

Socio-aware Public Safety Framework Design based on Contract Theory

3.1 Introduction

In public safety events, either natural disasters or terrorists attacks, the engagement of the citizens, the knowledge discovery, and the information dissemination play a critical role throughout the overall disaster management operation [67]. Nowadays, social networks have become of paramount importance in preparedness, emergency control management, response, and recovery. Millions of citizens depend on and exploit various social networks, such as Facebook, Weibo, Twitter, and others to spread information about critical events. This process in turn helps the Emergency Control Centers (ECC) to improve the disaster management operation. An indicative example is the "Boston Marathon" event in 2013, where the image of the suspect was retrieved from the social networks [68]. However, rumors and false information can also be disseminated in social networks and harm the rescue process in a public safety system [69]. Therefore, the classification, capabilities and interactions of the involved actors in the socio-aware public safety systems are of high significance.

Motivated by the aforementioned observations, in this chapter, we introduce a socio-aware public safety system, where the types of citizens offering information to the ECC are identified based on their social and communication characteristics. The ECC motivates the citizens to participate in the disaster management operation by offering to them incentives (e.g., benefits, coupons) in accordance to their types under information asymmetry or complete information, while the citizens contribute their personal effort to the process in order to improve the disaster management operation. To this end, we adopt *Contract Theory*, a powerful tool from microeconomics to model the citizens' incentive mechanism, through the use of contracts (agreements) between the ECC and the citizens. The ultimate goal is to find the optimal contract pair of ECC's offered reward and each citizen's provided effort based on its socio-communication profile and type.

3.1.1 Related Work

The actual and potential exploitation and impact of social networks on emergency disaster management and crisis situation has been studied in [70] to identify the benefits, e.g., monitoring situations, extending emergency response and management, as well as the negative developments, such as disseminating rumors. In [71], the authors introduce the concept of People as Sensors, where people contribute information through the social networks and their provided information is integrated within the location-based services, data analysis, and visualization systems. This concept is further extended in [72], where a tutorial of models and algorithms is presented for interactive sensing in social networks, where the users' provided information is exploited to optimize sensing, decision-making, and operations in dynamic environments, such as the public safety systems. Furthermore, a multimedia content analysis of the available information in social networks is introduced in [73] in order to detect events, e.g., natural disasters and terrorists' attacks, and manage the corresponding rescue operations.

Based on the above, it is evident that a great part of the available literature deals with the exploitation of the already available information in the social networks in order to detect public safety events and provide input to the disaster management operations. However, limited research has been performed in the area of properly modeling and exploiting the incentivization of the citizens in order to provide valuable information in the social networks that will support the ECC's operations. Towards this direction, some initial efforts have been devoted to encouraging citizens to report public safety problems, by capitalizing on the concepts of crowdsourcing, incentivization and volunteer computing [74, 75]. Nevertheless, the majority of them have been relatively primitive focusing primarily on finding ways of simply engaging citizens, being either heuristic or crude in their nature, without attempting to quantify the contribution of each citizen in a formal manner.

In this chapter, we adopt concepts and principles of *Contract Theory*, which provides the mathematical foundations to design formal and informal agreements to motivate people with potentially conflicting interests to take mutually beneficial actions, which otherwise would be counter-productive. Under this concept, an employer provides contracts to the employees based on their profiles, i.e., types, to motivate them to provide back their effort, which is crucial for the employer's operational processes.

Contract theory has been already applied in several communication-related applications, including device-to-device (D2D) communications [76] and cooperative spectrum sharing [77, 78]. In particular, in [79], the authors introduce a contract-theoretic relay selection framework, where the employer is the transmitter and the employees are the relay nodes. The transmitter offers rewards, i.e., payments, to the employees, while the latter guarantee a signal-to-interference-plus-noise-ratio at the destination. Contract theory has been also used to incentivize the users, i.e., employees, to establish device-to-device communication pairs to de-congest their communication with the base station (i.e., employer) [80]. Also, contract theory is used in cooperative spectrum sharing [81], and in cognitive networks allowing the primary spectrum owner (i.e., employer) to incentivize the secondary users (i.e., employees) to efficiently share the available bandwidth [82].

3.1.2 Contributions & Outline

This chapter aims exactly at filling the aforementioned research gap, by introducing formal methods - based on the Contract Theory - in order for the ECC to incentivize the citizens to participate in the disaster management operations and offer their valuable effort and information in an optimal manner. The key scientific contributions of our work that differentiate it from the rest of the existing literature, are summarized as follows.

1. The different types of the citizens are identified by the ECC via exploiting their social and communication characteristics, and identify a socio-communication type for each citizen. Based on the citizen's type, a corresponding utility is formulated reflecting its perceived satisfaction from the received reward for its invested effort. Also, the ECC's utility is defined to capture the overall benefit of using the citizens' efforts, while considering the corresponding cost of providing incentives to the citizens through the rewards. A contract pair between the ECC and each citizen is considered to be established consisting of the ECC's reward and the citizen's effort (Section 3.2).

2. The problem of determining the optimal rewards from the ECC's perspective and

the optimal invested effort from the citizens' side is formulated and solved initially considering that the ECC has complete information about the types of the citizens (Section 3.3). Furthermore, the aforementioned problem is addressed and thoroughly analyzed, under the most challenging and realistic assumption of ECC's incomplete information knowledge about the citizens' types (Section 3.4). In both scenarios, the outcome of the proposed framework is the optimal contract pairs.

3. A series of simulation experiments are realized to evaluate the performance and inherent attributes of the proposed socio-aware public safety framework (Section 3.5). Finally, Section 3.6 concludes the chapter.

3.2 System Model

We consider a public safety system consisting of an Emergency Control Center (ECC) that is responsible to coordinate the disaster management operations and a set of citizens $C = \{1, \ldots, c, \ldots, |C|\}$. The ECC rewards the citizens through personalized rewards r_c (e.g., benefits, coupons, money) in order to incentivize them to provide their valuable effort in the disaster management operations. The citizen's effort q_c can capture various types of effort: (a) social related effort, such as data quality (e.g., sensing data, closed cameras TV data), information shared in social networks, influential posts on Twitter, shelters' announcements on Facebook, and others, and (b) communication related effort, such as coverage area of the citizen's mobile device, which can potentially act as a relay node, CPU capability provided by the citizen's devices to process data in a fog computing setup and others. We consider the normalized values of the ECC's rewards, i.e., $r_c \in [0, 1]$, and the citizen's effort, i.e., $q_c \in [0, 1]$. Also, the ECC acts in a fair manner and rewards more the citizens that provide more effort in the disaster management operation, thus the reward r_c is a strictly increasing function with respect to the citizen's effort q_c.

3.2.1 Citizen's Social and Communication Type

Each citizen is characterized by a socio-communication type t_c which captures its social and communication characteristics in terms of providing information to the ECC. Regarding the communication characteristics, each citizen achieves a data rate $R_c = Wlog(1 + \frac{P_c g_c}{\sum_{i \neq c} P_i g_i + I_0})$ to directly report information to the ECC's receiver, where W is the system's bandwidth, P_c and g_c are the citizen's transmission power and channel gain, respectively, $\sum_{i \neq c} P_i g_i$ is the overall sensed interference, and I_0 is the background noise [83]. It is evident that the greater the citizen's achievable data rate is, the more valuable it becomes for the ECC's operation as more information can be collected by the ECC.

Moreover, the citizen's socio-communication type is also dependent on its social characteristics. Each citizen is characterized by its reputation score $\mu_c, \mu_c \in [0, 1]$, based on its activity in the social networks, i.e., information spread. The citizen's information spread is modeled in the literature by the diffusion model and the influence maximization algorithms can be used to determine the reputation score μ_c (i.e., identify the influential citizens) [84]. A citizen's contribution to the information spread process, is characterized by a corresponding social impact $SI_c(\mu_c)$ to the community, which is assumed a strictly increasing function with respect to the citizen's reputation score. Furthermore, to also capture the importance of the citizen's information contribution for the disaster management operation of the ECC, we introduce the concept of knowledge discovery $KD_c, KD_c \in [0, 1]$, referring to the unique content that the citizen shares to the social network or offers directly to the ECC compared to a bulk amount of data.

Based on the aforementioned citizen's social and communication characteristics, the socio-communication type $t_c, t_c \in [0, 1]$ of the citizen is defined as follows.

$$t_c = \frac{R_c}{\sum_{\forall i \in C} R_i} \cdot \frac{SI_c(\mu_c)}{\sum_{\forall i \in C} SI_i(\mu_i)} \cdot KD_c \tag{3.1}$$

For demonstration purposes, we also consider a strictly increasing function $r_c(q_c) = t_c q_c$ regarding the ECC's reward for citizen c. Also, in the following analysis, we consider $|C|$ different types of citizens, that is each citizen has a unique socio-communication type, while a citizen of higher type, i.e., $t_1 < \cdots < t_c < \cdots < t_{|C|}$, provides more effort q_c, i.e., $q_1 < \cdots < q_c < \cdots < q_{|C|}$.

3.2.2 Emergency Control Center's and Citizens' Utilities

The ECC offers a personalized contract pair $\{r_c(q_c), q_c\}$ to the citizen c for its provided effort q_c by providing a corresponding reward r_c. Each citizen is characterized by a utility function $U_c(q_c)$ expressing the perceived satisfaction from the ECC's provided reward based on its socio-communication type, as well as its cost to provide its effort that the citizen has invested in the disaster management operation. The citizen's utility is defined as follows.

$$U_c(q_c) = t_c \cdot e(r_c) - q_c \tag{3.2}$$

where $e(r_c)$ is the evaluation function of the citizen c regarding the received reward r_c. The evaluation function $e(r_c)$ is a strictly increasing, concave function with respect to the citizen's effort q_c, with $e(r_c = 0) = 0$ and expresses the citizen's satisfaction with respect to the reward that it received. For demonstration purposes and without loss of generality, in the following we consider $e(r_c) = \sqrt{r_c}$.

The ECC also experiences a utility $U^c_{ECC} = q_c - \kappa \cdot r_c$ by each citizen's provided effort, while taking into account the corresponding cost of the reward r_c (κ is the ECC's pricing factor). In the general case, the ECC may not be aware of the citizen's types, thus the ECC estimates them with probability p_c, where $\sum_{c=1}^{|C|} p_c = 1$. Thus, the ECC's overall perceived utility (accounting for all citizens) is defined as follows.

$$U_{ECC}(\mathbf{q}) = \sum_{c=1}^{|C|} [p_c(q_c - \kappa \cdot r_c)] \tag{3.3}$$

where $\mathbf{q} = (q_1, \ldots, q_{|C|})$ is the vector of the citizens' effort.

Considering the overall socio-aware public safety system, its social welfare, including both the ECC and all the citizens, is defined as follows.

$$SW(\mathbf{q}) = U_{ECC}(\mathbf{q}) + \sum_{c=1}^{|C|} U_c(q_c) \tag{3.4}$$

3.2.3 Contract Theory Perspective and Methodology

Based on the aforementioned utilities, rewards and other related parameters, in general the solution we seek is a set of contract pairs between the ECC and the citizens (employer and employees under Contract Theory terminology [39]) with reference to the citizen's effort q_c and the corresponding provided reward r_c, with the objective being maximizing the employer's utility. The problem is typically formulated as maximizing an objective function that represents the employer's utility, subject to the incentive compatibility constraint that the employee's expected utility is maximized when accepting the personalized contract, and the individual rationality constraint that the employee's utility under this contract is larger than or equal to its counterpart when not participating.

Contract theory is used to study the interaction between employer(s) and employees, and treats real world problems with either complete or even incomplete (often referred to as asymmetric information), by formally designing the contract between employer and employee, while implicitly introducing cooperation. The information asymmetry mainly refers to the fact that the employer does not know exactly the types and therefore the characteristics of the employees, and has only knowledge about the probability distribution of the types of employees. Contract-theoretic models allow the alleviation of this problem, and accordingly the employer can overcome this asymmetricity and efficiently still incentivize its employees. In the following sections we address the identification of optimal contract pairs between the ECC and citizens, for both cases of complete and incomplete information availability.

3.3 Citizens' Contracts under Complete Information

In this section, the ideal case where the ECC knows a priori the type of each citizen is considered. In this scenario which can be mainly used for benchmarking purposes, the ECC can fully exploit the citizens' efforts and make the best out of them regarding the disaster management operation. Thus, the ECC aims at maximizing its perceived utility by the effort of each citizen, while guaranteeing that the latter will accept the offered contract, i.e., the ECC has to ensure that the individual rationality condition of each citizen is satisfied. Therefore, the problem of determining the optimal contracts, under the assumption of complete information of the citizens' socio-communication types, can be written as follows.

$$\max_{\{r_c(q_c),q_c\}_{\forall c\in C}} U_{ECC}^{c} = q_c - \kappa \cdot r_c, \qquad \forall c \in C \tag{3.5a}$$

$$s.t. \qquad t_c \cdot e(r_c) - q_c \geq 0 \tag{3.5b}$$

The ECC will target at providing the minimum acceptable utility to the citizens towards maximizing its own utility. Thus, the constraint (3.5b) can be considered alternatively as equality in this case. Accordingly the solution of the optimization problem (3.5a)-(3.5b) is obtained by initially solving the equality (3.5b) with respect to r_c, and subsequently performing basic mathematical treatment and manipulations (i.e., substituting in (3.5a), differentiating Eq. (3.5a) with respect to q_c, and equating the outcome to zero). Consequently, under the assumption of complete information availability at the ECC, with respect to the exact type of each citizen and therefore its characteristics, the optimal contract pair is given by the following closed form solution: $\{r_c(q_c), q_c\} = \{(\frac{t_c}{2\kappa})^2, \frac{t_c^2}{2\kappa}\}$.

3.4 Contract Theoretic Public Safety Systems under Incomplete Information

In this section, we extend our study in determining the optimal contract pairs $\{r_c(q_c), q_c\}$ between the ECC and the citizens, under the realistic scenario of incomplete information availability, that is the ECC is not aware of the exact type of each citizen (i.e., information asymmetry). Nevertheless, the ECC should ensure two conditions for the citizens, i.e., individual rationality (IR) and incentive compatibility (IC), in order to guarantee their participation in the disaster management operation. The IR constraint refers to guaranteeing that the citizens will receive a non-negative utility by accepting the contract, thus, they will be at least willing to participate in the disaster management operation, while the IC constraint ensures that each citizen will receive the contract that better matches its type. The aforementioned conditions can be formally stated as follows.

Definition 1. *(Individual Rationality (IR)) A contract pair $\{r_c(q_c), q_c\}$ should guarantee that each citizen's utility is non-negative, i.e., $U_c(q_c) = t_c \cdot e(r_c) - q_c \geq 0, \forall c \in C$.*

Definition 2. *(Incentive Compatibility (IC)) Each citizen must select the contract pair $\{r_c(q_c), q_c\}$ designed for its type, i.e., $t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_{c'}) - q_{c'}, \forall c, c' \in C, c \neq$*

c'.

The IR and IC constraints are necessary, but not sufficient in order the ECC to determine the optimal contract pairs. Additionally, the following conditions and properties must hold true in order the contract pairs to be feasible.

Proposition 1. *For any feasible contract pair* $\{r_c(q_c), q_c\}$*, the following property must hold true:* $r_c > r_{c'} \Leftrightarrow t_c > t_{c'}$ *and* $r_c = r_{c'} \Leftrightarrow t_c = t_{c'}$.

Proof. Initially, we prove the sufficiency of the above property by using the IC constraint, i.e., $t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_{c'}) - q_{c'}, \forall c, c' \in C, c \neq c'$. Thus, we want to show $t_c > t_{c'} \Rightarrow r_c > r_{c'}$. Based on the IC constraint, we have:

$$t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_{c'}) - q_{c'} \tag{3.6}$$

$$t_{c'} \cdot e(r_{c'}) - q_{c'} \geq t_{c'} \cdot e(r_c) - q_c \tag{3.7}$$

By adding the inequalities (3.6) and (3.7), we have:

$$t_c e(r_c) + t_{c'} e(r_{c'}) \geq t_c e(r_{c'}) + t_{c'} e(r_c) \tag{3.8}$$

By continuing the derivations in inequality (3.8) and given that $t_c > t_{c'}$ and $e(r_c)$ is a strictly increasing function with respect to r_c, we conclude that $r_c > r_{c'}$. Continuing our analysis, we prove the necessity of the examined property, i.e., $r_c > r_{c'} \Rightarrow t_c > t_{c'}$. We have $r_c > r_{c'}$, and given that $e(r_c)$ is a strictly increasing function, we conclude that $e(r_c) - e(r_{c'}) > 0$. Based on Eq. 3.8, we have $t_c[e(r_c) - e(r_{c'})] \geq t_{c'}[e(r_c) - e(r_{c'})] \Leftrightarrow [t_c - t_{c'}][e(r_c) - e(r_{c'})] \geq 0$, thus $t_c > t_{c'}$. Similar analysis can be followed for the property $r_c = r_{c'} \Leftrightarrow t_c = t_{c'}$. □

The physical meaning of Proposition 1 is that a citizen of higher type t_c will receive a higher reward r_c compared to a citizen of lower type $t_{c'}$, who will receive lower reward $r_{c'}$. Proposition 1 guarantees the fairness in the rewards allocation from the ECC to the citizens.

Proposition 2. *(Monotonicity) A citizen of higher type, i.e.,* $t_1 < \cdots < t_c < \cdots < t_{|C|}$*, will receive a greater reward from the ECC, i.e.,* $r_1 < \cdots < r_c < \cdots < r_{|C|}$*, as it will contribute a greater effort, i.e.,* $q_1 < \cdots < q_c < \cdots < q_{|C|}$.

Proof. The proof of this proposition intuitively stems from Proposition 1, given that $t_1 < ... < t_c < ... < t_{|C|}$. □

In the following proposition, we examine the perceived utility of the citizens that have different socio-communication types.

Proposition 3. *A citizen of higher type, i.e.,* $t_1 < \cdots < t_c < \cdots < t_{|C|}$*, will receive a higher utility, i.e.,* $U_1 < \cdots < U_c < \cdots < U_{|C|}$.

Proof. We examine two citizens $c, c' \in C$ of types $t_c > t_{c'}, c \neq c'$. Based on the IC constraint, we have $t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_{c'}) - q_{c'} \xRightarrow{t_c > t_{c'}} U_c(q_c) = t_c \cdot e(r_c) - q_c > t_{c'} \cdot e(r_{c'}) - q_{c'} = U_{c'}(q_{c'})$. Thus, for $t_1 < \cdots < t_c < \cdots < t_{|C|}$, we conclude that $U_1 < \cdots < U_c < \cdots < U_{|C|}$. □

Based on the above introduced models, constraints, and the application of the key principles of Contract Theory, the ECC aims at maximizing its utility, while the citizens should satisfy all their personal constraints in order to be willing to participate in the socio-aware public safety system. Thus, optimization problem to determine the optimal contract pairs $\{r_c(q_c), q_c\}, \forall c \in C$ between the ECC and the citizens is formulated as follows.

$$\max_{\{r_c(q_c), q_c\}_{\forall c \in C}} U_{ECC}(\mathbf{q}) = \sum_{c=1}^{|C|} [p_c(q_c - \kappa \cdot r_c)] \tag{3.9a}$$

$$s.t. \qquad t_c \cdot e(r_c) - q_c \geq 0, \forall c \in C \tag{3.9b}$$

$$t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_{c'}) - q_{c'}, \forall c, c' \in C, c \neq c' \tag{3.9c}$$

$$0 \leq r_1 < \cdots < r_c < \cdots < r_{|C|} \tag{3.9d}$$

Given that the above optimization problem is non-convex, in the following we reduce its constraints in order to solve it in a tractable manner. Based on Proposition 2, we have $t_1 < \cdots < t_c < \cdots < t_{|C|}$ and considering the IC constraint, we have $t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_{c'}) - q_{c'} \geq t_c \cdot e(r_1) - q_1$. Given that $t_c > t_1$, we have: $t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_1) - q_1 \geq t_1 \cdot e(r_1) - q_1 \geq 0$. The last step of the latter inequality stems from the IR constraint (3.9b). Thus, if $t_1 \cdot e(r_1) - q_1 \geq 0$ holds true, then $t_c \cdot e(r_c) - q_c \geq 0$ holds true for each citizen $c \in C$. The above analysis concludes to the observation that if the IR constraint holds true for the citizen with the lower type, i.e., t_1, then it will hold true for any other citizen of higher type, thus, the IR constraints are reduced to $t_1 \cdot e(r_1) - q_1 = 0$. The latter IR constraint is considered as equality in order to the ECC to collect the maximum benefit from the citizen's effort.

In the following analysis, we target at reducing the IC constraints. The terminology that we use about the IC constraints between citizens: (a) $c, c', c' \in \{1, \ldots c-1\}$ is downward IC constraints, (b) $c, c', c' \in \{c+1, \ldots |C|\}$ is upward IC constraints, (c) $c, c-1, \forall c, c-1 \in C$ is local downward IC constraints, and (d) $c, c+1, \forall c, c+1 \in C$ is local upward IC constraints.

Proposition 4. *All the downward IC constraints can be represented by the local downward IC constraints.*

Proof. Considering three types of citizens: $t_{c-1} < t_c < t_{c+1}$, the local downward IC constraints can be written as:

$$t_{c+1} \cdot e(r_{c+1}) - q_{c+1} \geq t_{c+1} \cdot e(r_c) - q_c \tag{3.10}$$

$$t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_{c-1}) - q_{c-1} \tag{3.11}$$

Based on Proposition 1, we have $r_c > r_{c'} \Leftrightarrow t_c > t_{c'}$. For $r_c > r_{c-1} \overset{e\nearrow}{\Longleftrightarrow} e(r_c) > e(r_{c-1}) \Leftrightarrow e(r_c) - e(r_{c-1}) > 0$. Thus, for $t_{c+1} > t_c \Leftrightarrow t_{c+1}[e(r_c) - e(r_{c-1})] > t_c[e(r_c) - e(r_{c-1})] \geq^{(11)} q_c - q_{c-1}$. Therefore, we have recursively: $t_{c+1} \cdot e(r_{c+1}) - q_{c+1} \geq t_{c+1} \cdot e(r_{c-1}) - q_{c-1} \geq t_{c+1} \cdot e(r_{c-2}) - q_{c-2} \geq \cdots \geq t_{c+1} \cdot e(r_1) - q_1$. Thus, all the downward IC constraints can be equivalently captured by the local downward IC constraints:

$$t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_{c-1}) - q_{c-1} \tag{3.12}$$

□

Proposition 5. *All the upward IC constraints can be represented by the local downward IC constraints.*

Proof. Based on the IC constraint, we have:

$$t_{c-1} \cdot e(r_{c-1}) - q_{c-1} \geq t_{c-1} \cdot e(r_c) - q_c \tag{3.13}$$

$$t_c \cdot e(r_c) - q_c \geq t_c \cdot e(r_{c+1}) - q_{c+1} \tag{3.14}$$

Based on Proposition 1, we have $r_c > r_{c'} \Leftrightarrow t_c > t_{c'}$. Thus, based on Eq. 3.14, we have:

$$q_{c+1} - q_c \geq t_c[e(r_{c+1}) - e(r_c)] \geq t_{c-1}[e(r_{c+1}) - e(r_c)] \tag{3.15}$$

given that $t_c > t_{c-1}$. Based on Eq. 3.13, 3.15, we have: $t_{c-1}e(r_{c-1}) - q_{c-1} \geq t_{c-1}e(r_c) - q_c \geq t_{c-1}e(r_{c+1}) - q_{c+1}$. Thus, we have: $t_{c-1}e(r_{c-1}) - q_{c-1} \geq t_{c-1}e(r_{c+1}) - q_{c+1}$. Thus, if the IC constraint holds true for the citizen of type t_{c-1}, then all the upward IC constraints hold true. Therefore, we have recursively: $t_{c-1} \cdot e(r_{c-1}) - q_{c-1} \geq$

$t_{c-1} \cdot e(r_{c+1}) - q_{c+1} \geq \cdots \geq t_{c-1} \cdot e(r_{|C|}) - q_{|C|}$. Based on the above analysis, we conclude that the local upward IC constraints and all the upward IC constraints can be reduced to the local downward IC constraints. □

Based on the reduced IR constraints, and Propositions 4 and 5, the optimization problem (3.9a)-(3.9d) can be rewritten to the following convex optimization problem.

$$\max_{\{r_c(q_c), q_c\}_{\forall c \in C}} U_{ECC}(\mathbf{q}) = \sum_{c=1}^{|C|} [p_c(q_c - \kappa \cdot r_c)] \tag{3.16a}$$

$$s.t. \qquad t_1 \cdot e(r_1) - q_1 = 0, \forall c \in C \tag{3.16b}$$

$$t_c \cdot e(r_c) - q_c = t_c \cdot e(r_{c-1}) - q_{c-1} \tag{3.16c}$$

$$0 \leq r_1 < \cdots < r_c < \cdots < r_{|C|} \tag{3.16d}$$

The optimization problem (3.16a)-(3.16d) is solved using standard methods of convex optimization due to the convexity of the objective function and the constraints [85], and the optimal contract pairs $\{r_c(q_c), q_c\}$ are determined.

3.5 Numerical Results

In this section, a detailed numerical evaluation of the proposed contract-theoretic socio-aware public safety approach is conducted, via modeling and simulation. The performance evaluation initially focuses on the pure operation of the proposed framework in terms of determining the optimal contract pairs, the citizens' and the ECC's utilities, as well as the overall social welfare of the system, for both cases of complete and incomplete information availability (Section 3.5.1). Then, a comparative study of the proposed contract-theoretic framework against different alternative scenarios of determining the amount of effort offered by the citizens is presented (Section 3.5.2).

In the rest, we consider $\kappa = 0.999$, and that the probabilities of the citizens' types follow a uniform distribution. Moreover, the achievable data rate R_c is determined for each citizen considering a constant data transmission power $P_c = 2Watts$, the channel gain is $g_c = 1/d_c^2$, where $d_c \in [10, 400]m$ is the distance of the citizen c from the ECC's receiver, the system's bandwidth is $W = 5MHz$ [86], and the background noise is $I_0 = 10^{-13}$. The reputation score $\mu_c, \mu_c \in [0, 1]$ is appropriately calculated following the influence maximization algorithm [84, 87]. The social impact function is $SI(\mu_c) = log(\mu_c)$, and the knowledge discovery factor KD_c takes random values in the interval $[0, 1]$, where values closer to one indicate that the citizen has provided unique and valuable content to the ECC.

3.5.1 Pure Framework Operation Evaluation

Fig. 3.1.a presents the citizens' type values as a function of their index. We considered $|C| = 10$ indicative citizens, where the greater the citizen's index is the higher its type, i.e., $t_1 < \cdots < t_c < \cdots < t_{10}$. Fig. 3.1.b-3.1.d present the citizens' efforts, their offered rewards by the ECC and their achieved utilities as a function of the citizen's index, respectively, considering the scenarios of complete and incomplete information. Similarly, Fig. 3.1.e-3.1.f demonstrates the system's point of view, by presenting the ECC's utility and the system's social welfare, in a cumulative manner as the number of contributing citizens increases (each time inserting one additional citizen type indicated by the increased indices in the horizontal axis).

The results reveal that under the scenario of complete information (i.e., ideal scenario), the ECC knows a priori the socio-communication type of each citizen, thus,

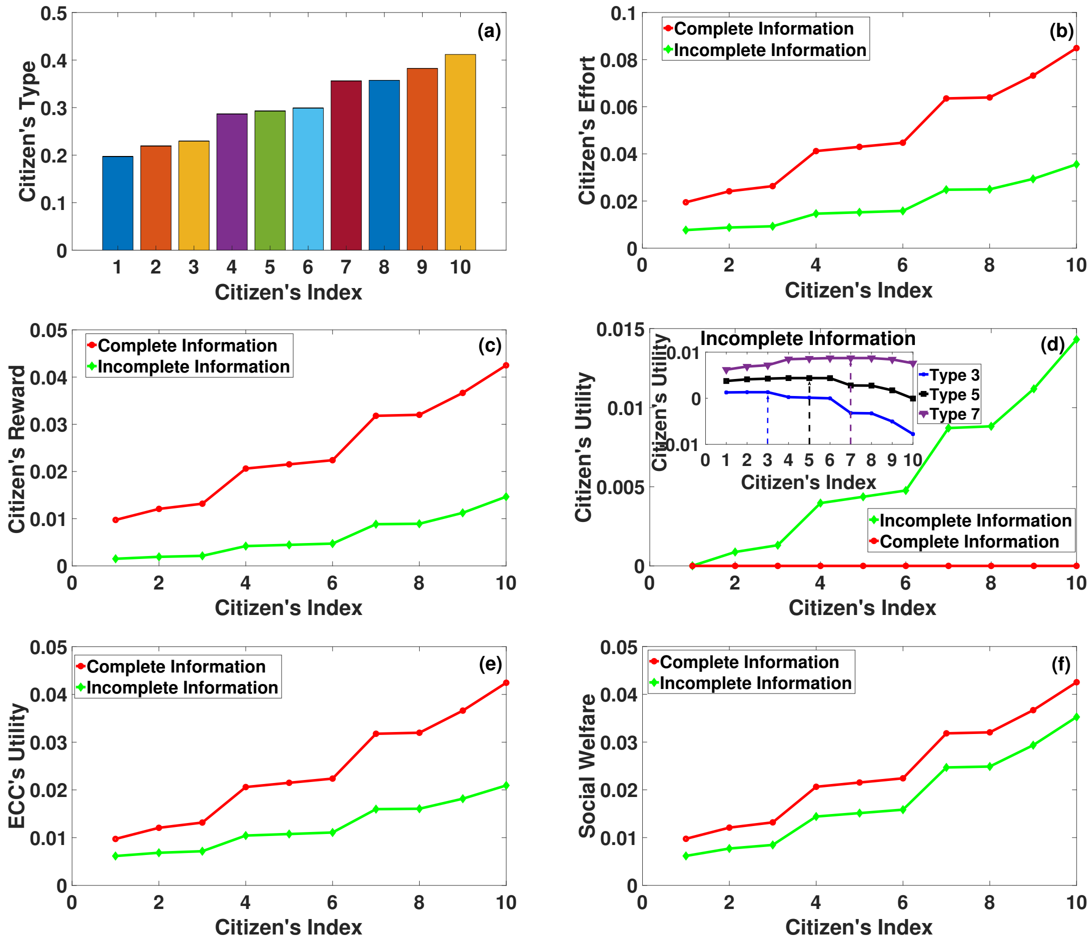

Figure 3.1: Pure Framework Evaluation - Complete and Incomplete (Asymmetric) Information Scenarios

it fully exploits the citizens' effort (Fig. 3.1.b) by providing increased rewards to them (Fig. 3.1.c), and achieving high ECC utility due to the increased citizens' participation (Fig. 3.1.e). Given that the ECC knows the citizens' types, it offers them the minimum possible reward based on their invested efforts in order to marginally satisfy their rationality constraints, thus $U_c = 0, \forall c \in C$ (Fig. 3.1.d).

On the other hand, under the incomplete information scenario, the ECC is not aware of the citizens' actual types, but it rather estimates them based on the knowledge about their probability distribution. In this case, the citizens by not disclosing their actual type to the ECC are able to achieve a higher utility compared to the complete information scenario (Fig. 3.1.d), i.e., tradeoff between their invested efforts (Fig. 3.1.b) and their rewards from the ECC (Fig. 3.1.c). Consequently, the ECC achieves lower utility compared to the complete information scenario (Fig. 3.1.e). The sub-graph in Fig. 3.1.d shows that citizens' of higher type receive higher utility and the contract that matches the citizen's type concludes to the best achieved utility.

In a nutshell, based on Fig. 3.1.a-3.1.d, it is confirmed that a citizen of higher type, invests more effort, receives more reward from the ECC, and consequently achieves greater utility, eeeeeas also stated in Proposition 2. Moreover, in Fig. 3.1.f, we observe that despite the fact that under the incomplete information scenario the ECC is not aware of the exact type of each citizen, the achieved overall public safety system's social welfare is reduced only by approximately 15% for the case of $|C| = 10$ citizens (this value becomes even smaller for larger populations), which indicates that the proposed framework behaves very well under the challenging and realistic asymmetric scenario.

3.5.2 Comparative Evaluation

Fig. 3.2.a-3.2.c compares the proposed incomplete information realistic contact-theoretic framework's achieved ECC's utility, citizens' utilities, and overall system's social welfare, respectively, against three alternative strategies with respect to the citizen's effort investment, as follows: (i) minimum effort, (ii) maximum effort, and (iii) a random amount of effort. The results reveal that under the proposed framework the citizens are able to achieve high utility, similar to the one achieved by their minimum personal effort strategy (Fig. 3.2.b). Also, as expected, the ECC achieves the maximum utility if all the citizens invest their maximum effort (Fig.

3.2.a). However, despite the fact that ECC achieves low utility under the proposed contract-theoretic approach due to the cost of increased provided rewards to the citizens (owing to the information incompleteness assumption), the system's social welfare is the highest among all scenarios (Fig. 3.2.c). The latter shows that the proposed framework enables the smooth collaboration between the ECC and the citizens, concluding to improved social welfare. Finally, the scenario where the citizens invest a random effort presents an intermediate trend regarding all the examined metrics, between the minimum and the maximum invested efforts scenarios.

In the following we compare the strategy where the ECC offers personalized rewards to the citizens (according to their type as realized in the proposed framework, i.e., $r_c = t_c \cdot q_c$), against an alternative still linear but type agnostic reward approach, offering common reward to all the citizens, i.e., $r_c = \frac{\sum_{c=1}^{|C|} t_c}{|C|} \cdot q_c$. We observe that the citizens benefit in terms of their achieved utility under the contract-theoretic (CT) framework, while the ECC achieves lower utility compared to the linear reward scenario, as in the latter case it tends to over-reward the citizens without adopting to their socio-communication type (Fig.3.3.a). We also observe that the contract-theoretic framework achieves higher system's social welfare (Fig. 3.3.b), exceeding by approximately 25% the corresponding values under the linear reward framework.

3.6 Conclusions

In this chapter, a socio-aware public safety framework founded on the properties of contract theory is proposed, in order to determine the optimal contract pairs between the ECC and the citizens, towards incentivizing the latter to participate in the disaster management operation. The citizens are characterized by their socio-communication type, capturing both their activity in the social networks and their communication characteristics. The citizens provide their efforts to the ECC, which in return rewards them. The identification of the optimal contract pairs, i.e., ECC's rewards and citizens' efforts, have been provided under the scenarios of both complete and incomplete information, with respect to the ECC knowledge about the actual type of each citizen.The overall framework was evaluated via modeling and simulation, in terms of its efficiency and effectiveness, by studying multiple operation approaches and scenarios. Part of our current and future work contains the extension of this model under the principles of Prospect Theory, towards capturing the citizens' behavioral characteristics in their utilities under risks and uncertainty, and their corresponding impact on the optimal contract pairing.

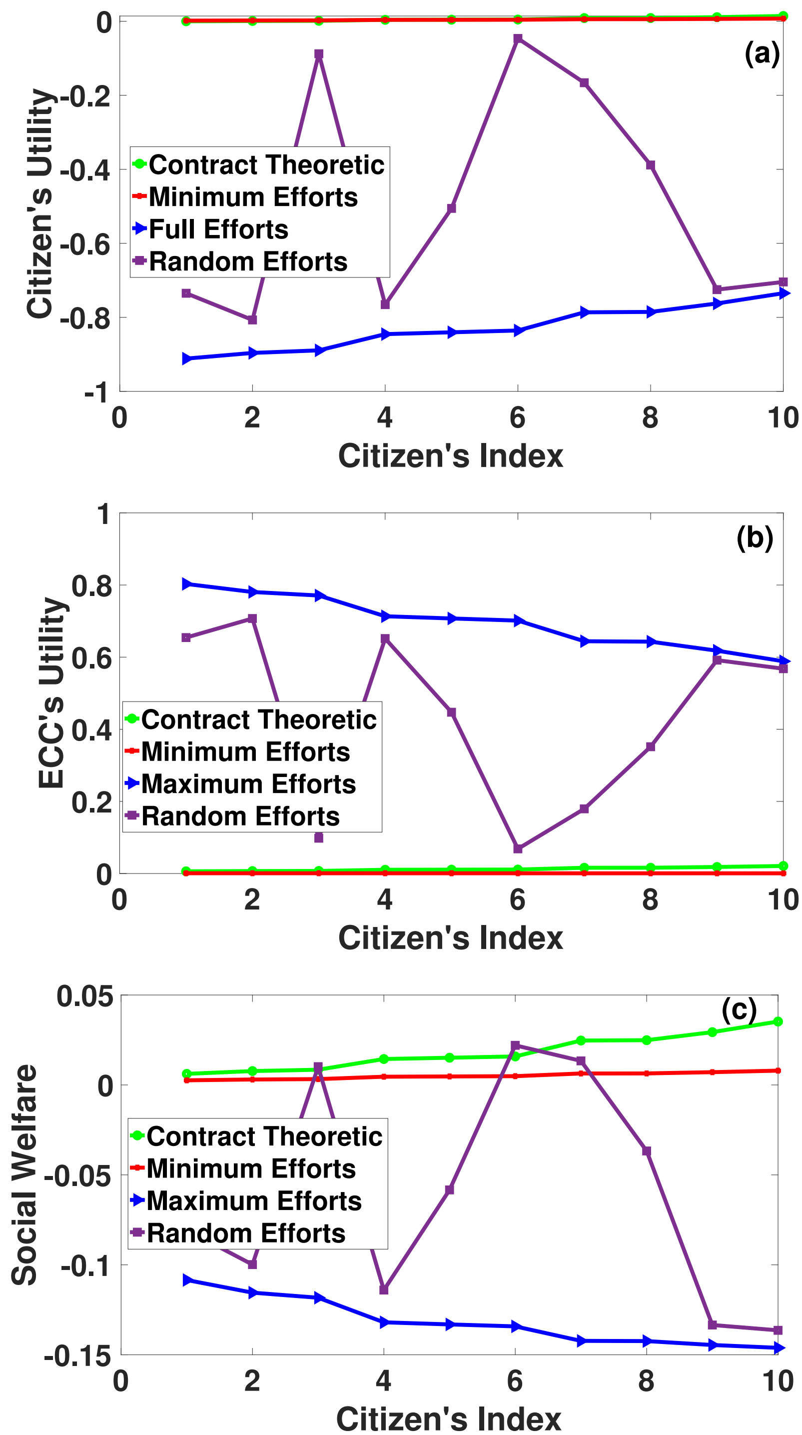

Figure 3.2: Comparative Evaluation

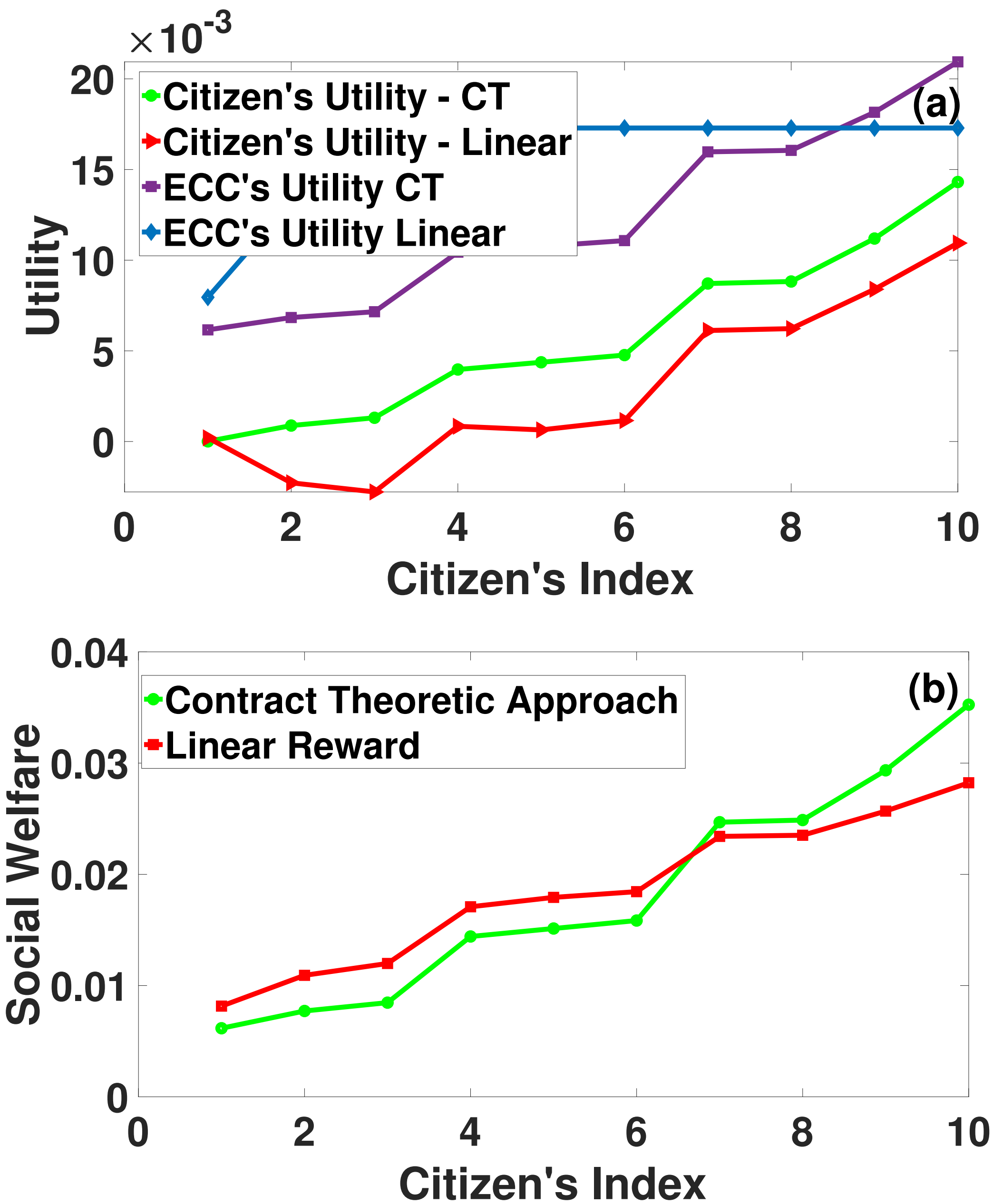

Figure 3.3: Type dependent vs. type agnostic rewards

Chapter 4

UAV-enabled Dynamic Multi-Target Tracking and Sensing Framework

4.1 Introduction

Unmanned Aerial Vehicles (UAVs) have attracted the interest of the research community due to their salient attributes, such as strong line-of-sight connection links, fast and flexible deployment and mobility. Their vital features have enabled them to support various civil Internet of Things (IoT) applications, such as surveillance systems [88]. UAVs have also been used for data collection from critical areas in crowdsourcing applications [89]. Motivated by these applications, in this chapter we propose a UAV-enabled multi-target tracking and sensing framework, where the UAVs are matched to the targets based on a reputation model, and the optimal data collection is determined in a distributed manner by a game-theoretic approach.

4.1.1 Related Work & Motivation

Computer vision-based target tracking is proposed in the literature using the sparse representation theory to model the target's appearance [90]. In [91], the target tracking problem is formulated based on the partially observable Markov decision process framework, where input is provided by an on board camera. The joint problem of target tracking and UAV path planning is studied in [92], by using vision sensors, a laser scanner, and an on board embedded computer. A deep reinforcement learning (DRL) approach is proposed in [93] to deal with the target tracking problem, under the challenge of frequent changes of the target's aspect ratio. In [94], the authors determine the minimum number of UAVs that are needed to detect a set of targets by formulating a network flow-based problem and solving it with heuristic algorithms.

UAVs have also been used to support crowdsourcing IoT applications enabling the data collection from targets residing in critical areas, e.g., public safety scenarios. In [95], a UAV-assisted crowd surveillance use case is studied, where the UAVs collect videos from cameras on the ground and they process them either on board or at the ground servers. In [96], the UAV's flight time is minimized by optimizing its altitude, while jointly maximizing the number of offloaded bits by the ground devices. In [97], the joint optimization problem of the UAV's trajectory and radio resource allocation is studied via a successive convex approximation framework, to maximize the number of served devices in terms of achievable uplink data rate.

However, despite the significant advances achieved by these efforts, they either neglect or partially consider, the problem of stable matching among the UAVs and the targets, as well as the incentivization of the targets to provide their data to the UAVs. In this chapter, we aim to address this research gap by introducing (i) a holistic reputation model to evaluate the targets' potential to provide useful data,

(ii) an intelligent matching framework between the UAVs and the targets, and (iii) a game-theoretic approach to determine the targets' optimal amount of offloaded data to the UAVs, while following a pricing-based approach to incentivize them to perform the data offloading.

4.1.2 Contributions & Outline

The key technical contributions of this research work are summarized as follows.

- A reputation model is introduced to quantify the targets' reputation in terms of valuable offloaded data to the UAVs. It consists of (i) the *UAV-agnostic reputation*, where the targets' reputation is determined by all the UAVs, and (ii) the *trustworthy reputation*, where the evaluation of a trusted set of UAVs regarding the targets' reputation weighs more (Section 4.2).

- Representative preference matching functions are formulated for the UAVs and the targets to capture their preferences in terms of pairing among each other. An intelligent matching algorithm is realized to decide the targets to be tracked by the UAVs (Section 4.3).

- The targets and the UAVs utility from offloading and collecting data, respectively, is captured in utility functions. A Stackelberg game is formulated among each target and the associated UAV to determine each target's optimal amount of offloaded data and the effort-based price that the UAV offers to the target to incentivize it to offload its data. The properties of existence, uniqueness and convergence to the Stackelberg Equilibrium are proven (Section 4.4).

- A set of detailed numerical results is presented to evaluate the performance of the proposed framework, while a comparative study demonstrates its superiority in terms of successful target tracking and data collection (Section 4.5).

4.2 Models & Assumptions

4.2.1 System Model

We consider a snapshot of a smart city environment consisting of a set of targets $I = \{1, \ldots, i, \ldots, |I|\}$ (e.g., ambulances, firetrucks, mobile IoT sensors), and a set of UAVs $N = \{1, \ldots, n, \ldots, |N|\}$. The position of each UAV at the time t is $p_t^n = (x_t^n, y_t^n, z_t^n)$. The target's position $q_t^i = (x_t^i, y_t^i, 0m)$ at time t is stochastic following a bivariate Gaussian distribution. Thus, the UAVs know the likelihood $\phi^i(q_t^i) : Q \to \mathbb{R}_{>0}$ that the target i is at a location q_t^i at time t. We obtain the highest likely probabilistic position $\hat{q}_t^i = (\hat{x}_t^i, \hat{y}_t^i, 0m)$ by employing the mean of the target's Gaussian distribution. Each UAV n is characterized by its normalized flying time $F_n \in [0, 1]$, which depends on its energy availability, where a value closer to one indicates a greater flying time. Each target i has a personal normalized cost $c_i \in (0, 1]$ (e.g., consumed energy) to collect the data $d_{i,n}$ that will be offloaded to a UAV n, thus, it charges the UAV with an effort-based price $P_{i,n}$ in order to obtain its data. For generalization purposes, we consider that the targets' data collection personal cost c_i and the effort-based price $P_{i,n}$ are unitless. Each target has a criticality factor $i_i \in (0, 1]$ based on the events in the surrounding environment. For example, an ambulance close to an area that a shooting occurred has greater criticality of data compared to a police car patrolling a neighborhood. The targets collect $D = \{1, \ldots, d, \ldots, |D|\}$ different types of data, e.g., videos, alerts, where $d \in (0, 1]$. A greater value of d represents an enhanced type of data, e.g., video, compared to a smaller value of d, which indicates a lower type of data, e.g., speed alert. The popularity of each type of data is captured by the Zipf distribution $Zipf(d) = \frac{z_1}{d^{z_2}}$, $z_1 > 0, 0 < z_2 < 1$.

4.2.2 UAV-agnostic Reputation Model

The UAVs track the targets and collect data from them in order to report them to a central entity, e.g., the Emergency Control Center (ECC) in a smart city. Each target is characterized by a reputation based on how helpful or not was the provided information. In the *UAV-agnostic reputation model*, all the UAVs evaluate the targets' reputation that they interact with, each one with equal weight. Towards the UAV n evaluating how helpful is the information collected by the target i, the following metric is introduced: $H_{i,n} = \frac{d_{i,n}}{P_{i,n}} \cdot Zipf(d)$. Its physical notion is that a UAV considers the provided data from target i helpful if the data collection process is cost-efficient (i.e., $\frac{d_{i,n}}{P_{i,n}}$) and the type of the collected data is of high popularity (i.e., $Zipf(d)$). Thus, a binary parameter represents if the collected data are helpful ($c^{\lambda}_{i,n} = 1$, if $H_{i,n} \geq H_{thr}$) or not ($c^{\lambda}_{i,n} = 0$, if $H_{i,n} < H_{thr}$) for the UAV n in the λ-th interaction with the target i, where $H_{thr} = \sum_{\forall i \forall n} H_{i,n}/|I|$.

The reputation of a target i, as it is evaluated by a UAV n, decreases as the most recent interaction time among them elapses, given that the UAV has not a recent evaluation regarding the target's data. A reputation decay function $\log_2(\frac{b}{T-t^{\lambda}_{i,n}} + 1)$ is introduced, where $t^{\lambda}_{i,n}$ is the time instance of the λ-th interaction among the UAV n and the target i, T is the time duration that we study the sytem, and $b > 0$ is the decay factor. After each UAV is associated with a target (Section 4.3), the UAV n provides a good $GR_{i,n} = \sum\limits_{\lambda=1}^{\lambda_{i,n}} c^{\lambda}_{i,n} \cdot \log_2(\frac{b}{T-t^{\lambda}_{i,n}} + 1)$ or a bad reputation $BR_{i,n} = \sum\limits_{\lambda=1}^{\lambda_{i,n}} (1 - c^{\lambda}_{i,n}) \cdot \log_2(\frac{b}{T-t^{\lambda}_{i,n}} + 1)$ for the target i that is associated with, where $\lambda_{i,n}$ is the number of interactions among the target i and the UAV n in the examined duration T. Thus, the overall UAV-agnostic reputation that target i receives from UAV n, considering both its good and bad reputation, is derived as $UAR_{i,n} = \mathbb{E}(beta(GR_{i,n} + 1, BR_{i,n} + 1)) = \frac{GR_{i,n}+1}{GR_{i,n}+BR_{i,n}+2}$.

4.2.3 Trustworthy Reputation Model

In contrast to the UAV-agnostic reputation, there may be UAVs that their evaluation weighs more, e.g., UAVs belonging to the ECC, in the reputation score of a target. Thus, we determine the most trusted UAV $\hat{n} = \underset{n' \in N}{\text{argmin}}[\sum_{\substack{n \in N \\ n' \neq n}} |UAR_{i,n'} - UAR_{i,n}|]$ as the one that has the smallest difference from all the other UAVs for a specific target i. A UAV belongs to the set of trusted UAVs $N_{tr.,i}$ for a target i, if $|UAR_{i,\hat{n}} - UAR_{i,n}| \leq Tr_{thr}$, where $Tr_{thr} > 0$ is a trust threshold defined by the central entity.

The overall reputation of a target i combines the UAV-agnostic reputation and the trustworthy reputation. Thus, the overall good (Eq. 4.1) and the overall bad reputation (Eq. 4.2) of the target i is determined as follows.

$$OGR_{i,n} = w_1 \cdot GR_{i,n} + w_2 \cdot \sum_{n'=1}^{|N_{tr.,i}|} GR_{i,n'} \tag{4.1}$$

$$OBR_{i,n} = w_1 \cdot BR_{i,n} + w_2 \cdot \sum_{n'=1}^{|N_{tr.,i}|} BR_{i,n'} \tag{4.2}$$

where $w_1, w_2 \geq 0$ are the weighting factors of the UAV-agnostic and trustworthy reputation.

Thus, the overall reputation of the target i based on the evaluation of the UAV n is determined below.

$$R_{i,n} = \mathbb{E}(beta(OGR_{i,n} + 1, OBR_{i,n} + 1)) = \frac{OGR_{i,n} + 1}{OGR_{i,n} + OBR_{i,n} + 2} \tag{4.3}$$

4.3 Intelligent Multi-Target Tracking

In this section, an intelligent matching mechanism is introduced to pair each UAV with a corresponding target, while considering their tracking and sensing characteristics. Each UAV n has a preference function $M^t_{n,i}$ that captures its priority to track a target i in time t.

$$M^t_{n,i} = \frac{1}{|\hat{q}^i_t - p^n_t|} \cdot \frac{i_i}{P_{i,n}} \cdot \frac{R_{i,n}}{\sum\limits_{i \in I} R_{i,n}} \tag{4.4}$$

The physical notion of Eq. 4.4 is that a UAV prefers to track a target that is in its close proximity, has high criticality of collected data, provides its data in a competitive effort-based price, and it has a good reputation.

Each target i has a preference function $TM_{i,n}$ that captures its priority to offload data to a UAV n at time t.

$$TM^t_{i,n} = \frac{1}{|\hat{q}^i_t - p^n_t|} \cdot \frac{F_n}{c_i} \cdot \frac{R_{i,n}}{|UAR_{i,\hat{n}} - UAR_{i,n}|} \tag{4.5}$$

The physical notion of Eq. 4.5 is that a target i prefers to offload its data to a UAV n that (i) is in its close proximity, thus the target will spend less energy to offload the data; (ii) has a long flying time, thus the target has sufficient time to transmit its data; (iii) the target's data collection cost for the requested amount of data by the UAV n is low; and (iv) is trustworthy and has provided an overall high reputation for the target i.

Based on the above, we build the UAVs' and the targets' matching tables at time t, as $M^t = (M^t_{i,n})_{|I|\times|I|}$ and $TM^t = (TM^t_{i,n})_{|N|\times|N|}$, respectively. We consider $|N| = |I|$, and we are searching for a stable matching among the UAVs and the targets by examining the problem from the UAVs' perspective. Following the matching theory, we adopt the Gale-Shapley algorithm [98] to enable the UAVs to select the targets that will track at every examined time t. The main steps of the proposed multi-target multi-UAV matching algorithm are as follows.

1. At each time t, the UAVs and the targets have ranked the members of the opposite set based on their own preference function, i.e., Eq. 4.4 and Eq. 4.5, respectively.

2. Each UAV, which is not already paired with a target, will be randomly chosen to propose to its most preferable target (as indicated by the UAV's matching table M^t), which has not already rejected this UAV.

3. The target being proposed will: (i) accept the UAV's proposal, if this is the target's first received proposal; (ii) reject if this proposal is worse (in terms of the target's preference order of UAVs) than its current proposal; and (iii) accept if this proposal is better than its current one.

4. If all the UAVs are paired, the matching algorithm stops, otherwise returns to step 2.

The outcome of the multi-target multi-UAV matching algorithm is the stable pairs (i^*, n^*) of targets and UAVs.

4.4 Optimal Sensing

In this section, the problem of optimal sensing, i.e., data collection from the smart city's field, is addressed. Given the pairs of UAVs and targets, the target's i utility by offloading $d_{i,n}$ data to the UAV n, is given as follows.

$$U_{i,n}(P_{i,n}, d_{i,n}) = P_{i,n} \cdot d_{i,n} - c_i \cdot d_{i,n} \tag{4.6}$$

where $c_i = \frac{k_i}{Zipf(d)}, k_i > 0$ is a personalized cost (e.g., energy cost) of the target i to collect the data of type d. The target's utility represents the revenue $(P_{i,n} \cdot d_{i,n})$ that the target gains by offloading its data, while considering its corresponding cost $(c_i \cdot d_{i,n})$ to collect the data.

The experienced utility of a UAV n by tracking a target i and collecting data from it, is formulated as follows.

$$\mathcal{U}_{n,i}(P_{i,n}, d_{i,n}) = \mu_n \cdot \log_2(1 + \sum_{i \in I} R_{i,n} d_{i,n}) - \sum_{i \in I} P_{i,n} d_{i,n} \tag{4.7}$$

where $\mu_n > 0$ is the UAV's n operation factor, i.e., level of contribution to the smart city's proper operation. It is noted that the UAVs belong to a central entity of the smart city, that controls the data collection operation. The first term of Eq. 4.7 represents the perceived utility of the UAV n by the available information in the smart city field that is collected by the targets. The second term of Eq. 4.7 represents the smart city central entity's total cost (charged by the targets) to collect the data.

Each target aims at maximizing its utility during the data collection process by determining the optimal effort-based price $P^*_{i,n}$ that will charge the UAV in order to provide its data $d_{i,n}$. Each target's utility maximization problem is formulated as follows.

$$\max_{P_{i,n}} U_{i,n}(P_{i,n}, d_{i,n}) \tag{4.8}$$

Similarly, each UAV aims at maximizing its own utility during the data sensing operation. Each UAV determines the optimal amount of data $d^*_{i,n}$ that it can receive from the target that is paired with, while providing the corresponding effort-based price. Each UAV's utility maximization problem is formulated as follows.

$$\max_{d_{i,n}} \mathcal{U}_{n,i}(P_{i,n}, d_{i,n}) \tag{4.9}$$

The two utility maximization problems of the target (Eq. 4.8) and the UAV (Eq. 4.9) are coupled together through the variables $P_{i,n}$ and $d_{i,n}$. Thus, we follow a two-step *Stackelberg game-theoretic approach*, where the target i is the leader and the UAV n is the follower. The Stackelberg game is played between a UAV n and a target i, thus, $|I|$ Stackelberg games are played in parallel at time t. Towards determining the Stackelberg Equilibrium (SE) of each game, we perform a backward induction.

The UAV determines its optimal sensing demand of data $d^*_{i,n}$ requested from the target towards maximizing its utility, as follows: $\frac{\partial \mathcal{U}_{n,i}}{\partial d_{i,n}} = \frac{\mu_n R_{i,n}}{1+\sum\limits_{i \in I} R_{i,n} d_{i,n}} - P_{i,n}$ and $\frac{\partial^2 \mathcal{U}_{n,i}}{\partial d^2_{i,n}} = -\frac{\mu_n R^2_{i,n}}{(1+\sum\limits_{i \in I} R_{i,n} d_{i,n})^2} < 0$. We observe that $\mathcal{U}_{n,i}$ is strictly concave with respect to the requested amount of data $d_{i,n}$. Thus, it has a unique optimal amount of data $d^*_{i,n}$ determined as follows.

$$d^*_{i,n} = [\frac{\mu_n}{P_{i,n}} - \frac{1 + \sum\limits_{i' \in I, i' \neq i} R_{i',n} d_{i',n}}{R_{i,n}}]^+ \tag{4.10}$$

where $[x]^+, x \geq 0$. Based on Eq. 4.10, we derive the following observations: (i) the sensing demand of data $d_{i,n}$ of the UAV n is proportional to the target's i overall reputation and inversely proportional to the target's i effort-based price that charges the UAV; (ii) the targets compete with each other to gain a higher reputation by reducing the effort-based price, thus, reducing their personal cost.

The target's utility function (Eq. 4.6) can be rewritten as $U_{i,n}(P_{i,n}, d^*_{i,n}) = (P_{i,n} - c_i) \cdot [\frac{\mu_n}{P_{i,n}} - \frac{1+\sum\limits_{i' \in I, i' \neq i} R_{i',n} d_{i',n}}{R_{i,n}}]$, based on Eq. 4.10. It is noted that if the effort-based price $P_{i,n}$ that a target i charges a UAV n is high, this will impact the UAV's tracking decision (Section 4.3), and the UAV may select another target to track. Thus, the target's optimal effort-based price $P^*_{i,n}$ is the Best Response to the other targets announced prices, i.e., $P^*_{i,n} = BR(\mathbf{P_{-i,n}})$, where $\mathbf{P_{-i,n}} = (P_{1,n}, \ldots, P_{i-1,n}, \ldots, P_{i+1,n}, \ldots, P_{|I|,n})$. Towards proving the existence and uniqueness of an SE, we show that the target's utility function is strictly concave with respect to the effort-based price $P_{i,n}$, as follows: $\frac{\partial U_{n,i}}{\partial d_{i,n}} = \frac{\mu_n c_i}{P^2_{i,n}} - \frac{1+\sum\limits_{i' \in I, i' \neq i} R_{i',n} d_{i',n}}{R_{i,n}}$ and $\frac{\partial^2 U_{n,i}}{\partial d^2_{i,n}} = -\frac{2\mu_n c_i}{(P_{i,n})^3} < 0$. Thus, the best response strategy of the target i is:

$$P^*_{i,n} = BR(\mathbf{P_{-i,n}}) = \sqrt{\frac{R_{i,n} \mu_n c_i}{1 + \sum\limits_{i' \in I, i' \neq i} R_{i',n} d_{i',n}}} \tag{4.11}$$

Based on Eq. 4.10, 4.11, the SE is $(P^*_{i,n}, d^*_{i,n})$ for the Stackelberg game played among the UAV n and the target i. In order to prove the convergence of the target's i best response strategy to the SE, it suffices to prove that $P^*_{i,n} = BR(\mathbf{P_{-i,n}})$ is a standard function [99, 100].

Theorem 1. *Each target's $i, i \in I$, best response strategy $BR(\mathbf{P_{-i,n}})$ is a standard function.*

Proof. Towards proving Theorem 1, the properties of positivity, monotonicity, and scalability should hold true.

1. *Positivity:* Based on Eq. 4.11, we have $BR(\mathbf{P_{-i,n}}) > 0$.

2. *Monotonicity:* Base on Eq. 4.10, 4.11, we have $BR(\mathbf{P_{-i,n}}) = \sqrt{\frac{R_{i,n}\mu_n c_i}{1+\sum\limits_{i' \in I, i' \neq i} R_{i',n}[\frac{\mu_n}{P_{i',n}} - \frac{1+\sum\limits_{i'' \in I, i'' \neq i} R_{i'',n} d_{i'',n}}{R_{i',n}}]}}$. Thus, we observe that $P_{i',n}$ is proportional to $BR(\mathbf{P_{-i,n}})$. Therefore, the property of monotonicity is satisfied.

3. *Scalability:* The following property should hold true: $a \cdot BR(\mathbf{P_{-i,n}}) > BR(a \cdot \mathbf{P_{-i,n}}), a > 1$. We have: $\frac{a \cdot BR(\mathbf{P_{-i,n}})}{BR(a \cdot \mathbf{P_{-i,n}})} = \sqrt{\frac{a^2 + \sum\limits_{i' \in I, i' \neq i} aR_{i',n}[\frac{\mu_n}{P_{i',n}} - \frac{1+\sum\limits_{i'' \in I, i'' \neq i} R_{i'',n} d_{i'',n}}{R_{i',n}}]}{\sum\limits_{i' \in I, i' \neq i} R_{i',n}[\frac{\mu_n}{P_{i',n}} - \frac{1+\sum\limits_{i'' \in I, i'' \neq i} R_{i'',n} d_{i'',n}}{R_{i',n}}]}}$.

Given that $a > 1$, we have $\frac{a \cdot BR(\mathbf{P_{-i,n}})}{BR(a \cdot \mathbf{P_{-i,n}})} > 1 \iff a \cdot BR(\mathbf{P_{-i,n}}) > BR(a \cdot \mathbf{P_{-i,n}})$. Thus, we conclude that $P^*_{i,n} = BR(\mathbf{P_{-i,n}})$ is a standard function with respect to $\mathbf{P_{-i,n}}$. □

4.5 Numerical Results

In this section, a detailed numerical evaluation is presented in terms of (i) the proposed reputation model's success to capture the system's conditions (Section 4.5.1); (ii) the performance of the intelligent matching algorithm (Section 4.5.2); (iii) the

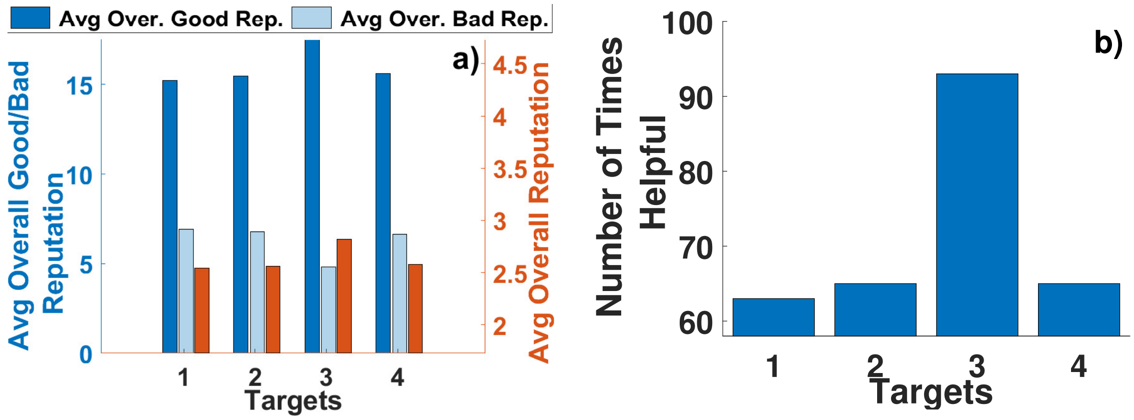

Figure 4.1: Targets reputation model – Targets perspective

operation of the game-theoretic sensing framework (Section 4.5.3); and (iv) the benefits of the overall framework compared to other alternatives (Section 4.5.4). For the purposes of the evaluation, the values of the considered key parameters are as follows: $|N| = |I| = 4, b = 0.5, Tr_{thr} = 0.1, w_1 = 0.6, w_2 = 0.4, z_1 = d^*_{i,n}, z_2 = 1/P^*_{i,n}$, an area of $100m \times 100m, z^n_t = 121m, T = 100, \mu_n = [1.115, 1.355, 1.675, 1.789]$, while F_n, i_i randomly distributed in $(0, 1]$. The proposed framework's evaluation was conducted in a HP Laptop, 1.8GHz Intel Core i7, with 16GB LPDDR3 available RAM.

4.5.1 Operation of Targets Reputation Model

In the following we examine the operation of the reputation model, both from the targets and the UAVs perspective. In particular, initially Fig. 4.1a presents the targets' average overall good (Eq. 4.1), overall bad (Eq. 4.2), and overall (combined) reputation (Eq. 4.3), over the time period of $T = 100$ time instances, while Fig. 4.1b depicts the number of times that the targets were providing helpful data to their associated UAVs. The results confirm that the targets with the highest average overall good reputation and the smallest average overall bad reputation conclude to better average overall reputation (Fig. 4.1a). Accordingly, as shown in Fig. 4.1b their provided data to the UAVs are evaluated as helpful more times.

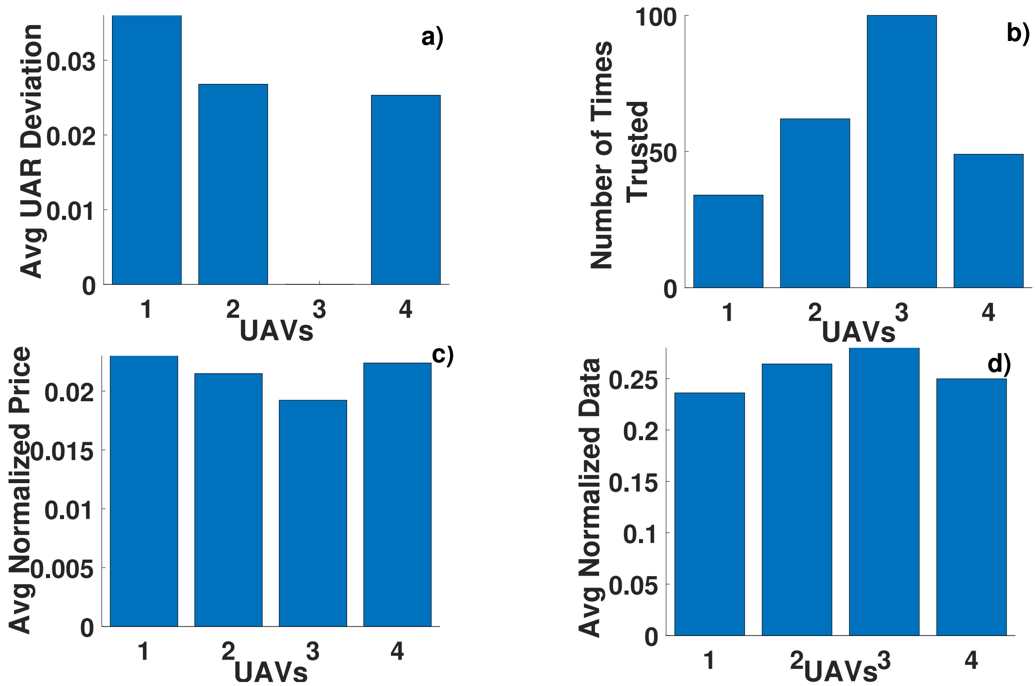

Figure 4.2: Targets reputation model – UAVs perspective

Towards examining the operation of the proposed reputation model from the UAVs' perspective, Fig. 4.2a-4.2d present the UAVs' agnostic reputation deviation from the most trusted UAV, i.e., $|UAR_{i,\hat{n}} - UAR_{i,n}|$, the number of times that each UAV belongs to the set of trusted UAVs $N_{tr.,i}$ of its associated target, the average normalized effort-based price that it experiences and the average normalized amount of data that it collects, respectively. We observe that the UAVs with the smallest deviation (Fig. 4.2a) are trusted more times (Fig. 4.2b). Thus, based on the outcome of the SE of each game among each UAV and its associated target, they collect more data (Fig. 4.2d) by investing a smaller effort-based price (Fig. 4.2c), thus collectively concluding to more cost-efficient data sensing.

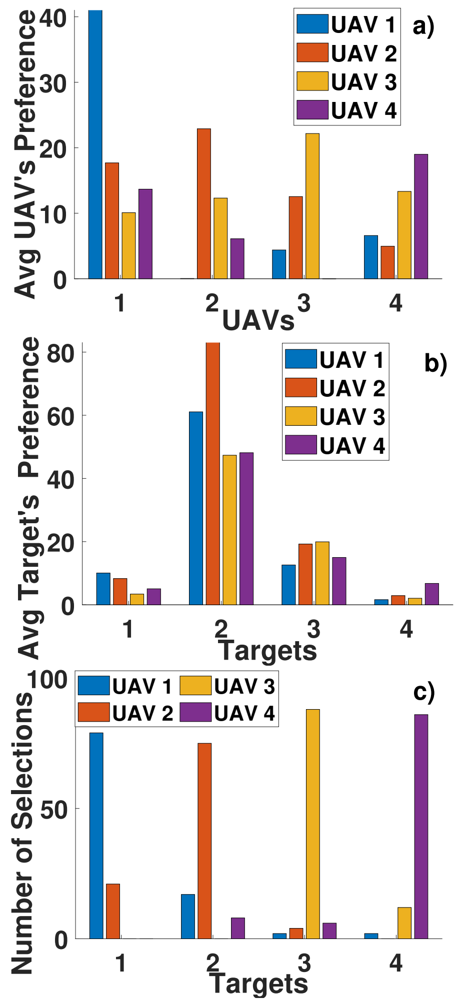

Figure 4.3: UAVs – Targets intelligent matching

4.5.2 UAVs – Targets Intelligent Matching Framework

The following results in Fig. 4.3a-4.3c demonstrate the operation and effectiveness of the introduced UAVs-targets matching framework, in terms of the UAVs' preferences (Eq. 4.4), the targets' preferences (Eq. 4.5), and the actual number of targets' selections by the UAVs for a time duration $T = 100$ time instances, respectively. Specifically, based on the results illustrated in Fig. 4.3a, it is observed that UAV 1 prefers to track target 1, UAV 2 prefers to track target 2, etc. The exact symmetric observation holds true regarding the targets' preferences to offload their data to the

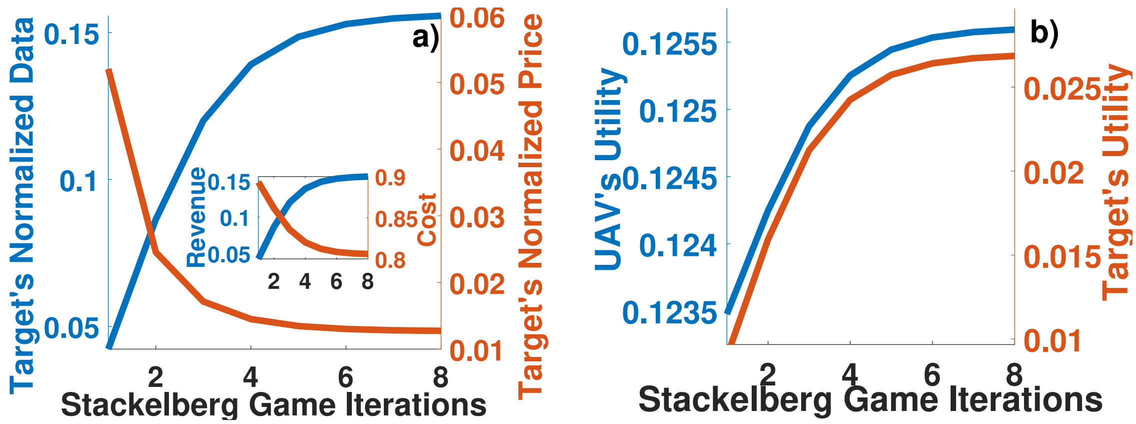

Figure 4.4: Optimal sensing – Convergence to the SE

corresponding UAVs (Fig. 4.3b). It is noted that the proposed matching framework captures in a holistic manner both the UAVs and the targets matching preferences through the proposed preference functions, i.e., Eq. 4.4, 4.5, thus concluding to an overall successful matching outcome (Fig. 4.3c).

4.5.3 Optimal Sensing Framework Operation Evaluation

In the following, the operation of the optimal sensing framework (Section 4.4) is evaluated, and the convergence of the corresponding game to the unique SE is shown. The Stackelberg game between one UAV and the target that is associated with, is examined for one time instance t. Fig. 4.4a present the target's normalized offloaded data $d_{i,n}$ and the corresponding effort-based price $P_{i,n}$, as a function of the game's iterations. The enclosed subfigure presents the respective target's revenue and cost. Fig. 4.4b presents the target's utility $U_{i,n}$, and the UAV's utility $\mathcal{U}_{n,i}$ as a function of the number of iterations. The results reveal that the target's offloaded amount of data and the corresponding price (Fig. 4.4a) converge monotonically to the SE in few iterations (less than 8 iterations equivalent to 7msec). Following the outcome of the Stackelberg game, the target's and the UAV's utility (Fig. 4.4b) also monotonically converge to the optimal outcome given the uniqueness of the SE. Also, during the

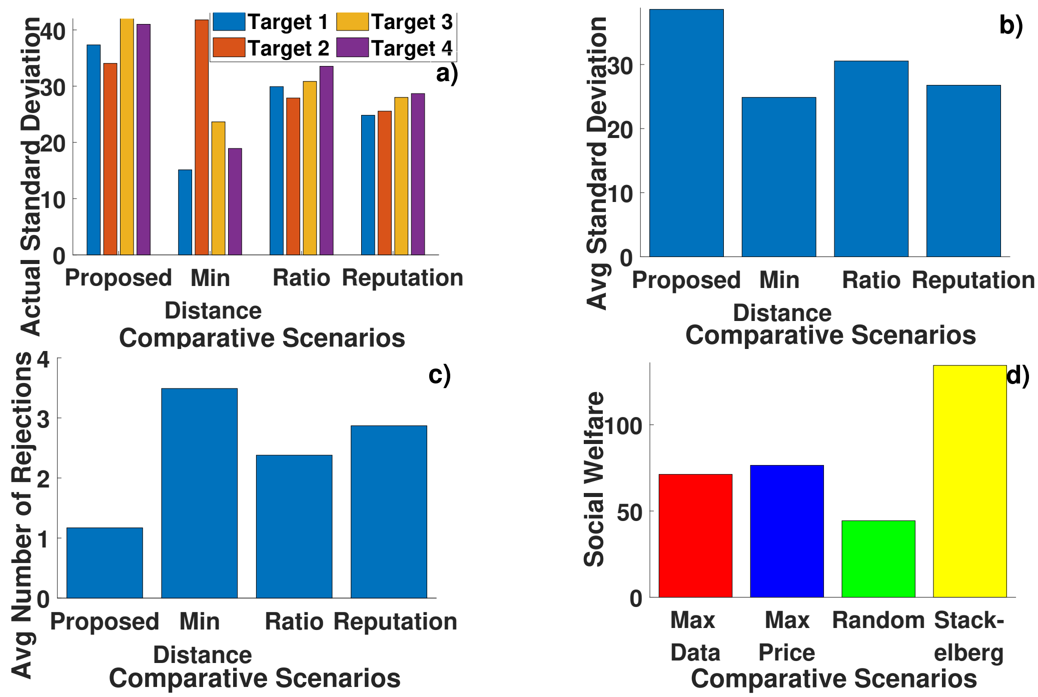

Figure 4.5: Comparative Evaluation

Stackleberg game's iterations, the target increases its revenue and decreases its cost by strategically deciding its offloaded data, while considering the effort-based pricing limitations (Fig. 4.4a).

4.5.4 Comparative Evaluation

In this section, initially we compare the proposed intelligent matching framework with the following three alternative matching approaches. (1) Ratio Approach: The UAVs select the targets that have high criticality of collected data and provide their data in a competitive effort-based price, i.e., $M^t_{n,i} = \frac{i_i}{P_{i,n}}$. The targets select the UAVs that have long flying time and its personal cost to collect the data is low, i.e., $TM^t_{i,n} = \frac{F_n}{c_i}$. (2) Reputation Approach: The UAVs and the targets define their preferences based on the reputation model, i.e., $M^t_{n,i} \frac{R_{i,n}}{\sum\limits_{i \in I} R_{i,n}}$, $TM^t_{i,n} = \frac{R_{i,n}}{|UAR_{i,\hat{n}} - UAR_{i,n}|}$.

(3) Min Distance Approach: The UAVs' and the targets' preferences are defined based on the minimum distance between them, i.e., $M^t_{n,i} = TM^t_{i,n} = \frac{1}{|\hat{q}^i_t - p^n_t|}$.

Fig. 4.5a and Fig. 4.5b present the actual standard deviation of the number of selections of each UAV from the most selected target and the corresponding average standard deviation over all the targets in the system for all the comparative approaches, respectively. Furthermore, Fig. 4.5c presents the corresponding average number of rejections, i.e., two UAVs preferred the same target and due to conflict one UAV's preference was rejected. The results reveal that under the proposed matching framework, the UAVs experience few conflicts among each other (Fig. 4.5c), while they tend to select their most preferable target, and therefore the actual (Fig. 4.5a) and average (Fig. 4.5b) standard deviation of the number of selections from their most preferable selection is high. The Ratio approach presents also small number of conflicts among the UAVs (Fig. 4.5c) compared to the Reputation and the Min Distance approaches, due to the great variation of the UAVs' preference function given the personalized price $P_{i,n}$ that target i charges UAV n. In the Reputation approach, all the UAVs tend to select the most reputable targets, while in the Min Distance approach, the closest targets. Thus, in those two approaches, the number of rejections is high (Fig. 4.5c) and the actual (Fig. 4.5a) and average (Fig. 4.5b) standard deviation of the number of selections from their most preferable selection are consequently low.

Additionally, we compare the proposed optimal sensing framework against the following three alternatives: (1) Max Data Scenario: All targets offload their total amount of collected data. (2) Max Price Scenario: The targets charge the UAVs with a fixed (maximum) price. (3) Random Scenario: The targets decide randomly the amount of data to offload and the price to charge. For fairness purposes, in all comparative approaches, the intelligent matching algorithm introduced in this chapter, is adopted. The social welfare of the system, i.e., the summation of the targets' (Eq. 4.6) and the UAVs' utilities (Eq. 4.7), is presented in Fig. 4.5d for all the considered comparative scenarios for $T = 100$ time instances. The results clearly reveal the superiority of the proposed optimal sensing framework, while the Max Data and the Max Price scenarios both present similar low social welfare levels, and the Random approach provides the worst outcome.

4.6 Conclusions

In this chapter, a novel holistic UAV-enabled multi-target tracking and sensing framework is introduced. Initially, each target's reputation is defined, consisting of both UAV-agnostic and trustworthy reputation models. Based on that, the intelligent pairing of the UAVs with the targets towards enabling the multi-target tracking by the UAVs, is performed. The targets' optimal data offloading strategies along with the optimal effort-based price that the UAVs are charged with in order to collect the targets' data, are determined based on a Stackelberg game-theoretic approach. Detailed numerical results were presented highlighting the key operational features and the performance benefits of our proposed approach. Part of our current and future work focuses on treating the examined problem based on a labor economics approach under the principles of Contract Theory, towards incentivizing the targets to offload their data to the UAVs [101].

Chapter 5

Contract-Theoretic Resource Control in Public Safety Systems

5.1 Introduction

Public Safety Networks (PSNs) have been introduced to provide reliable exchange of data during catastrophic events (e.g., natural disasters, terrorist attacks). The persistent and robust information flow in disaster-struck areas has been enabled by the usage of Unmanned Aerial Vehicles (UAVs). UAV-enabled wireless communications have attracted great research and commercial interest due to their salient attributes, i.e., controllable mobility, line-of-sight communication with the transmitters, and low-cost, fast, and flexible deployment [102]. Moreover, the Wireless Powered Communications (WPC) networking paradigm enables the mobile devices to harvest energy from the radio frequency signals of the transmitter [103, 104]. Capitalizing on the advances achieved by these technologies, in this chapter, we consider a UAV-assisted WPC network that enables the efficient data collection from a disaster-struck area, following a contract-theoretic approach.

5.1.1 Related Work & Motivation

The problem of maximizing the system's energy efficiency in a three layer UAV-assisted network architecture (space-air-ground) is studied in [105] considering an Internet of Remote Things network, where the UAVs act as relays. The authors formulate and solve an optimization problem to determine the devices' subchannel selection, their optimal transmission power, and the UAVs' deployment. In [106], a UAV performs the data collection from an Internet of Things (IoT) field. The authors jointly optimize the UAV's flying speed, altitude, and the IoT devices' frame length at the MAC layer, to maximize the ground sensors energy efficiency. In [107], an ant colony optimization algorithm is presented that enables the collaboration between the UAVs and the ground devices, in order to prolong the lifetime of the network, by reducing the devices' energy consumption to report their data to the UAVs.

The concept of UAV-enabled WPC system has been introduced in [108], where UAV-mounted energy transmitters, transmit radio frequency signals and the ground devices harvest energy from them. In [109], the UAV's trajectory is obtained to maximize the harvested energy by the ground devices under the UAV's flying speed and altitude constraints. In [110], the authors aim at maximizing the minimum achievable throughput of the ground devices, by jointly optimizing the UAVs' trajectories, the users' transmission power, and the decision between the devices' energy harvesting and information transmission phases.

However, all these research efforts have been conducted in isolation focusing on only one of the following related problems, that is: the energy efficient information acquisition from the ground nodes, the energy harvesting from the UAVs' radio frequency signals, and/or the optimal UAVs' deployment. This fragmentation has not yet allowed the exploitation of the corresponding achievements in their full capacity. Accordingly, in this chapter, we aim to address this research gap by introducing an

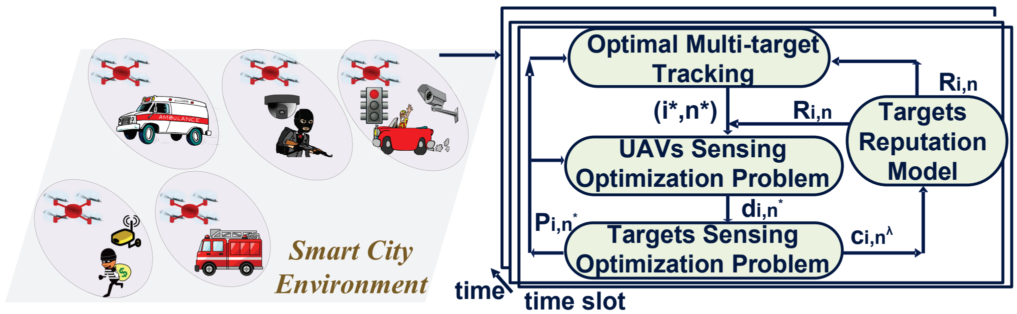

Figure 5.1: UAV-assisted WPCN topology and framework's architecture

energy efficient information flow and energy harvesting framework capturing users' risk-aware characteristics, based on the principles of *Contract Theory*, and the support of *Reinforcement Learning*.

5.1.2 Contributions & Outline

The main contributions of this research work are summarized as follows.

- A wireless powered communication network (WPCN) assisted by UAVs [111] charging the victims' devices is considered in a public safety scenario [112]. The victims' risk-aware characteristics to provide their information to the UAVs are captured in representative utility functions [113]. An optimization problem determining each victim's optimal amount of provided information to the UAV and each UAV's optimal charging power, is formulated and solved following the principles of Contract Theory, by introducing a labor market relationship among the UAVs and the users (Section 5.2).

- The victims are organized in rescue groups and the rescue leaders are determined for each group through a socio-physical groups formation mechanism (Section 5.3.1). A reinforcement learning framework, based on the theory of *Stochastic Learning Automata*, is introduced to enable the optimal matching between the UAVs and the rescue leaders of each group, in a distributed and efficient manner (Section 5.3.2).

- A set of simulation experiments are performed demonstrating the basic characteristics of the proposed contract-theoretic framework, while considering users' risk-aware behavior. The benefits of the proposed framework are highlighted in terms of energy-efficiency, information acquisition from the disaster area, and intelligent users' incentivization to support the rescue operation (Section 5.4).

5.2 Contract-theoretic Control of Resources

A UAV-assisted WPCN is considered within a public safety system consisting of a set of victims $V = \{1, \ldots, v, \ldots, |V|\}$, a set of UAVs $U = \{1, \ldots, u, \ldots, |U|\}$, and the Emergency Control Center (ECC). The channel gain between two victims v, v' is defined as $G_{v,v'} = \frac{\lambda}{d^2_{v,v'}}$, where $\lambda > 0$ represents the channel fading and $d_{v,v'}$ [m] is the distance among the victims v and v' [114, 115]. Let E_v [J] denote the energy availability of each victim's v device and d_v [m] represent the distance of the victim from the source of the disaster (e.g., epicenter of an earthquake). The victims are organized in rescue groups. Each rescue group rg determines its rescue leader rl_{rg} following a socio-physical rescue groups formation mechanism (Section 5.3.1). Each rescue leader selects in a distributed manner to which UAV it will offload its data based on a reinforcement learning approach (Section 5.3.2). The considered system's topology is presented in Fig. 5.1. Initially, we assume that the rescue groups formation and the rescue leaders association to the UAVs have already been performed and we focus on the contract-theoretic control of the resources.

During a catastrophic event, the ECC needs to collect information from the victims in order to plan the rescue operation [116]. Thus, incentives should be offered to them in order to provide information to the UAVs and correspondingly to the

ECC. At the same time, the victims' behavioral characteristics, i.e., risk-aware behavior in terms of providing information, should be considered, while designing their incentives. To achieve this goal, the principles of Contract Theory are adopted [39]. Contract Theory is a powerful tool to design effective incentives by modeling the UAVs-victims relation based on a labor market setup. Specifically, the victims of a rescue group report their information to the corresponding rescue leader. Then, a UAV, which collects information from the rescue leader, considers the rescue leader's risk averse characteristics and offers rewards (i.e., incentives) in order to compensate it for its invested labor (i.e., reporting information).

Each victim transmits with power proportional to the normalized distance from its rescue leader, i.e., $P_v = \frac{d_{v,rl_{rg}}}{\max\limits_{v \in V_{rg}} d_{v,rl_{rg}}} \cdot P_v^{max}$, where P_v^{max} is the victim's maximum transmission power and V_{rg} is the set of victims belonging to the rescue group rg. The corresponding achievable transmission data rate is $R_v = W \cdot \log(1+\frac{G_{v,rl_{rg}}P_v}{\sum\limits_{v' \geq v+1}^{|V'|} G_{v',rl_{rg}}P_{v'}+I_0})$ [117], where I_0 represents the Additive White Gaussian Noise and W [Hz] is the system's bandwidth, while non-orthogonal multiple access has been considered, and the successive interference cancellation technique is implemented at the receiver, i.e., rescue leader. Thus, during a timeslot t [sec], the total amount of data that the rescue leader collects is: $\mathbb{D}_{rl_{rg}} = (\sum\limits_{v \in V_{rg}} R_v)t$ [bits].

Each UAV offers a contract to each rescue leader that is associated with. The contract is defined as $(w_{rl_{rg}} \cdot P_u^{max}, TD_{rl_{rg}})$, where P_u^{max} is the UAV's maximum charging power and $TD_{rl_{rg}}$ are the collected data from its rescue group, where $TD_{rl_{rg}} \leq \mathbb{D}_{rl_{rg}}$. We consider the UAV's provided reward as $w_{rl_{rg}} \in [0,1]$, thus, the corresponding charging power is $w_{rl_{rg}} \cdot P_u^{max}$. Each rescue leader invests an effort (i.e., labor) $a_{rl_{rg}} \in [0,1]$ and transmits $TD_{rl_{rg}} = a_{rl_{rg}} \cdot \mathbb{D}_{rl_{rg}}$ data to the UAV. The rescue leader's performance, as it is evaluated by the UAV, is defined as $q_{rl_{rg}} = a_{rl_{rg}} + \epsilon$, where ϵ represents some noisy data. The parameter ϵ follows a normal distribution with zero mean and variance σ^2. Towards capturing the rescue leader's risk aware characteristics in terms of reporting information to the UAV, as well as its perceived satisfaction from its action and the harvested energy from the UAV, the rescue leader's risk aware utility function is defined as follows [39].

$$U_{rl_{rg}}(w_{rl_{rg}}, a_{rl_{rg}}) = -e^{-n_{rl_{rg}}[w_{rl_{rg}} - \psi(a_{rl_{rg}})]} \tag{5.1}$$

where $n_{rl_{rg}} \in (0, 1]$ is the rescue leader's risk aversion parameter. The greater the value of $n_{rl_{rg}}$ is, the more conservative the rescue leader becomes in terms of uploading information to the UAV in order to save its own energy. The function $\psi(a_{rl_{rg}})$ is the cost function of the rescue leader capturing its personal cost (energy consumption) to report the collected information from the disaster area to the UAV. The cost function is concave with respect to the rescue leader's invested effort, e.g., $\psi(a_{rl_{rg}}) = \frac{ca^2_{rl_{rg}}}{2}$, where $c > 0$ is a constant cost factor. The reward percentage $w_{rl_{rg}}$ offered by the UAV is defined as $w_{rl_{rg}} = \mu + s_{rl_{rg}} \cdot q_{rl_{rg}}$, where μ is a fixed compensation level, i.e., $\mu \cdot P_u^{max}$, to reward the rescue leaders for even participating in the information flow process, and $s_{rl_{rg}}$ is the variable compensation related to the rescue leader's performance component. The contract-theoretic control problem of the UAVs (i.e., charging power) and the rescue leaders (i.e., transmitted data) resources is formulated as a maximization problem of the UAV's expected profit.

$$\max_{a_{rl_{rg}}, s_{rl_{rg}}} \mathbb{E}(q_{rl_{rg}} - w_{rl_{rg}}) \tag{5.2a}$$

$$\mathbb{E}(-e^{-n_{rl_{rg}}[w_{rl_{rg}} - \psi(a_{rl_{rg}})]}) \geq U_{rl_{rg}}|_{min} \tag{5.2b}$$

$$a_{rl_{rg}} \in \underset{a_{rl_{rg}}}{\operatorname{argmax}}\, \mathbb{E}(-e^{-n_{rl_{rg}}[w_{rl_{rg}} - \psi(a_{rl_{rg}})]}) \tag{5.2c}$$

where $U_{rl_{rg}}|_{min}$ is the minimum acceptable utility by the rescue leader in order to be motivated to send the collected data. The constraint (5.2b) represents the *individual rationality* constraint of the rescue leader. If this inequality does not hold true, then, the rescue leader has no incentive to report the collected data to the UAV.

The constraint (5.2c) captures the *incentive compatibility* for each victim, i.e., each victim will put an effort to report the collected data in order to maximize its own perceived utility.

The rescue leader's expected utility can be written as $\mathbb{E}(-e^{-n_{rl_{rg}}[w_{rl_{rg}}-\psi(a_{rl_{rg}})]}) = -e^{-n_{rl_{rg}}[\mu+s_{rl_{rg}}a_{rl_{rg}}-\frac{ca^2_{rl_{rg}}}{2}-\frac{n_{rl_{rg}}s^2_{rl_{rg}}\sigma^2}{2}]})$ given that we can show that $\mathbb{E}(-e^{-n_{rl_{rg}}s_{rl_{rg}}\epsilon}) = e^{\frac{n_{rl_{rg}}s^2_{rl_{rg}}\sigma^2}{2}}$ from the theory of the normal distribution. Thus, by solving the constraint (5.2c), we can determine the rescue leader's optimal amount of transmitted data to the UAV.

$$TD^*_{rl_{rg}} = a^*_{rl_{rg}} \cdot \mathbb{D}_{rl_{rg}} = \frac{s_{rl_{rg}}}{c} \cdot \mathbb{D}_{rl_{rg}} \tag{5.3}$$

We can eliminate the constraint (5.2c) by substituting Eq. 5.3 to Eq. 5.2a and rewrite the optimization problem.

$$\max_{a_{rl_{rg}}, s_{rl_{rg}}} [\frac{s_{rl_{rg}}}{c} - (\mu + \frac{s^2_{rl_{rg}}}{c})] \tag{5.4a}$$

$$s.t. \quad \mu + \frac{s^2_{rl_{rg}}}{c} - \frac{c}{2}\frac{s^2_{rl_{rg}}}{c} - \frac{n_{rl_{rg}}}{2}\sigma^2 s^2_{rl_{rg}} = w_{rl_{rg}} \tag{5.4b}$$

The solution of the optimization problem (5.4a, 5.4b) yields to the optimal UAV's reward, i.e., charging power.

$$w^*_{rl_{rg}} \cdot P_u^{max} = [\mu + \frac{1}{1+n_{rl_{rg}}c\sigma^2}(\frac{s_{rl_{rg}}}{c} + \epsilon)]P_u^{max} \tag{5.5}$$

Thus, the optimal contract among a UAV and rescue leader is determined to be $(w^*_{rl_{rg}} \cdot P_u^{max}, TD^*_{rl_{rg}})$. The operational timeslot of the system is splitted into the wireless energy transfer (WET) phase with duration τ_h[sec] and the wireless information transmission (WIT) phase with duration τ_t [sec]. During the WET phase, the UAVs transfer directed energy to the rescue leaders that they are associated with, by unicasting a radio frequency signal via directional antennas [118]. The rescue leader's device's harvested energy from the UAV that it is associated with is given as follows.

$$HE_{rl_{rg}} = Eff_{rl_{rg}} \cdot \tau_h \cdot w_{rl_{rg}} * \cdot P_u^{max} \cdot G_{rl_{rg},u} \tag{5.6}$$

where $Eff_{rl_{rg}} \in (0,1]$ is the energy conversion efficiency factor, which depends on the rescue leader's device.

During the WIT phase, each rescue leader reports $TD^*_{rl_{rg}}$ to the UAV, assuming that its available energy, i.e., $E_{rl_{rg}} + HE_{rl_{rg}}$, is sufficient to report the contract theoretic optimal amount of data. Each rescue leader reports its optimal amount of data $TD^*_{rl_{rg}}$ through a dedicated subchannel with bandwidth W [Hz] to the UAV via adopting the single carrier frequency division multiple access (SC-FDMA) technique [119, 120, 121, 122]. Thus, its available data rate is $W \cdot \log(1 + \frac{G_{rl_{rg},UAV} P^{tr}_{rl_{rg}}}{I_0})$, where $P^{tr}_{rl_{rg}}$ is the rescue leader's transmission power. Thus, the rescue leader's consumed energy to transmit the $TD^*_{rl_{rg}}$ data is $E^{tr}_{rl_{rg}} = P^{tr}_{rl_{rg}} \cdot \tau_t$, and its remaining energy for the next timeslot is $E^{(t+1)}_{rl_{rg}} = E^{(t)}_{rl_{rg}} + HE_{rl_{rg}} - E^{tr}_{rl_{rg}}$.

5.3 Groups Formation and UAV Associations

5.3.1 Socio-physical Rescue Groups Formation

In this section, a socio-physical-aware rescue groups formation mechanism is presented, in order to enable the victims to create rescue groups and support the energy efficient information flow from the victims to the UAVs. In each rescue group, the victims transmit their information to the rescue leader of the group, who forwards it along with its own information to a UAV.

(1) *Physical Ties:* To support the victims' energy efficient communication, the victims tend to participate in rescue groups, where their communication distance among each other is small and their channel gain conditions are good. Thus, we define a symmetric matrix $G = \{g_{v,v'}\}_{|V|\times|V|}$, where $g_{v,v'} = \frac{G_{v,v'}}{\max\limits_{\forall v,v' \in V}\{G_{v,v'}\}} \in [0,1]$, which represents the normalized channel gain conditions of a pair of victims v, v'. Also, the victim's normalized energy availability $EA_v = \frac{E_v}{\max\limits_{\forall v' \in V}\{E_{v'}\}} \in [0,1]$ is critical in order to identify whether it could act as a rescue leader. The rescue leaders collect, process, and transmit the rest of the rescue group's victims' information, thus, they spend an increased amount of energy. Moreover, the victim's normalized distance from the source of the disaster $D_v = \frac{d_v}{\max\limits_{\forall v' \in V}\{d_{v'}\}} \in [0,1]$ is considered, as this victim can provide more accurate information to the UAV.

(2) *Social Ties:* The victims have interest to communicate with specific people, e.g., family members. The symmetric matrix $CI = \{ci_{v,v'}\}_{|V|\times|V|}, ci_{v,v'} \in [0,1]$ captures the victims' communication interest. A lower value of $ci_{v,v'}$ represents less communication interest among the victims.

By combining the victims' social and physical ties, we define a metric that captures the rescue and communication capability (RCC) of each victim, as follows.

$$RCC_v = EA_v \cdot D_v \cdot \sum_{v,v' \in V, v \neq v'} (ci_{v,v'} \cdot g_{v,v'}) \tag{5.7}$$

The socio-physical rescue groups formation mechanism is executed at the ECC, which informs the victims through broadcasted messages, and consists of the following steps.

(1) Initially, all the victims $|V|$ create a rescue group rg, whose set of victims is $V' = V$.

(2) For this rescue group rg with set of victims V', the rescue leader rl_{rg} is determined as $rl_{rg} = \underset{v \in V'}{\mathrm{argmax}}\{RCC_v\}$.

(3) For the victims that belong to the rescue group rg with set of victims V', the following condition must hold true:

$$g_{v,rl_{rg}} \cdot ci_{v,rl_{rg}} \cdot D_v \geq RG_{thres}^{(V')} \tag{5.8}$$

where $RG_{thres}^{(V')} = \frac{\sum\limits_{v \in V'} g_{v,rl_{rg}}}{|V'|} \cdot \frac{\sum\limits_{v \in V'} ci_{v,rl_{rg}}}{|V'|} \cdot \frac{\sum\limits_{v \in V'} D_v}{|V'|}$ is a threshold value to create homogeneous rescue groups in terms of consisting of victims with close distance, good channel conditions, high communication interest among each other, as well as contributing valuable information due to their proximity to the source of the disaster. The victims, who do not satisfy the condition (5.8), they form a new rescue group, with set of victims $V'' \subseteq V'$.

(4) Set $V' = V' - V''$ and if $|V'| > 1$, return to step 2, otherwise stop.

5.3.2 Reinforcement Learning-enabled Matching

In this section, a reinforcement learning-based framework is introduced to enable the optimal matching among the UAVs and the rescue leaders in a distributed and computationally efficient manner. Each leader acts as a stochastic learning automaton (SLA) making decisions of selecting a UAV to offload its data. Each UAV is characterized by a reputation, which depends on the physical and communication characteristics of the overall examined public safety system, and it is given as follows.

$$\mathcal{R}_u = \frac{\dfrac{\sum\limits_{rl_{rg} \in V_u} w_{rl_{rg}}^{*(ite-1)} P_u^{max} \sum\limits_{\forall rl_{rg}} d_{rl_{rg},u} \sum\limits_{rl_{rg} \in V_u} TD_{rl_{rg}}^{*(ite-1)} FT_u R_u}{\sum\limits_{\forall rl_{rg}} w_{rl_{rg}}^{*(ite-1)} P_u^{max} \sum\limits_{rl_{rg} \in V_u} d_{rl_{rg},u} \sum\limits_{\forall rl_{rg}} TD_{rl_{rg}}^{*(ite-1)} \sum\limits_{\forall u} R_u}}{|V_u^{(ite-1)}|^3} \tag{5.9}$$

where $FT_u \in (0,1)$ and R_u are the normalized flying time and the communications coverage radius of the UAV u, respectively, and V_u is the set of rescue leaders being served by the UAV u. The physical notion of Eq. 5.9 is that a rescue leader prefers to offload its data $TD^{*}_{rl_{rg}}$ to a UAV u that (a) collectively charges with high transmission power the rescue leaders that are connected to it; (b) is in close proximity; (c) has a long flying time and large communications coverage radius; (d) it tends to collect large amount of data; and (e) is not overcongested by other rescue leaders trying to simultaneously offload their data.

The probability of a rescue leader selecting the same UAV u to offload its data $TD^{*}_{rl_{rg}}$ in the next iteration of the SLA algorithm is given by Eq. 5.10a and the probability of selecting a different UAV is given by Eq. 5.10b [123].

$$Pr^{(ite+1)}_{rl_{rg},u} = Pr^{(ite)}_{rl_{rg},u} + b\mathcal{R}^{(ite)}_{u}(1 - Pr^{(ite)}_{rl_{rg},u}), u^{(ite+1)}_{rl_{rg}} = u^{(ite)}_{rl_{rg}} \tag{5.10a}$$

$$Pr^{(ite+1)}_{rl_{rg},u} = Pr^{(ite)}_{rl_{rg},u} - b\mathcal{R}^{(ite)}_{u} Pr^{(ite)}_{rl_{rg},u}, u^{(ite+1)}_{rl_{rg}} \neq u^{(ite)}_{rl_{rg}} \tag{5.10b}$$

where $u^{(ite)}_{rl_{rg}}$ is the selected UAV u by the rescue leader rl_{rg} in the iteration ite of the SLA algorithm and $0 < b < 1$ is the learning parameter that controls how fast the rescue leaders learn their optimal UAV matching. It is noted that the UAVs' reputation values are broadcasted by them to the rescue leaders to enable the latter execute the SLA algorithm in a distributed manner and eliminate the signaling overhead. The SLA algorithm converges when $Pr^{(ite)}_{rl_{rg},u} \geq Pr_{thr}$, $\forall rl_{rg}$ where Pr_{thr} is a threshold value, which for the evaluation purposes in this chapter is $Pr_{thr} = 0.95$. Then, each rescue leader offloads its data $TD^{*}_{rl_{rg}}$ to the selected UAV, as shown in Fig. 5.1.

5.4 Numerical Results

A detailed numerical evaluation illustrates the performance of the proposed framework in terms of the: impact of socio-physical parameters (Section 5.4.1), contract-

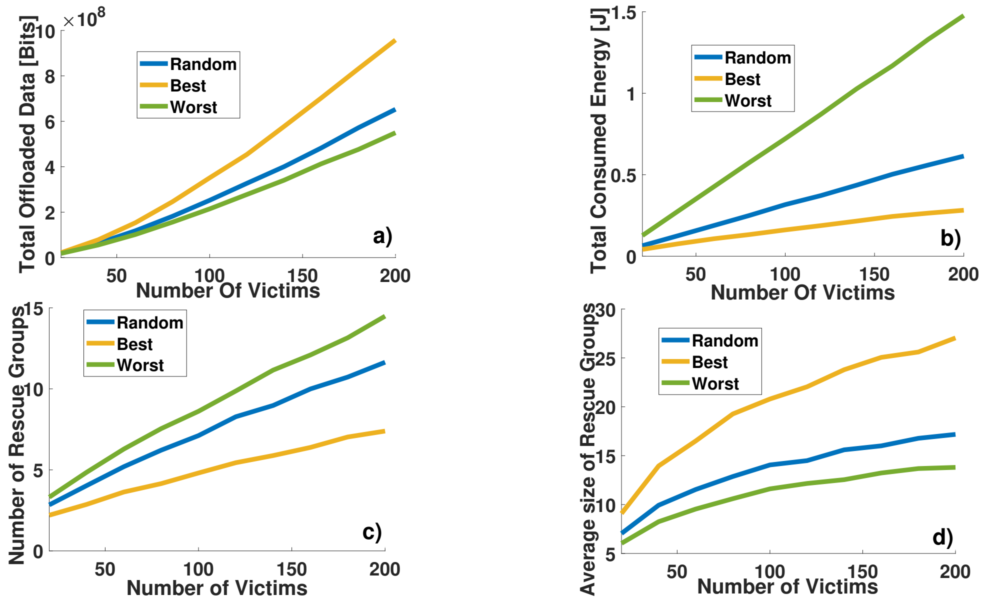

Figure 5.2: Impact of socio-physical parameters under different comparative scenarios

theoretic and behavioral-aware resource control (Section 5.4.2), and benefits of reinforcement learning to implement the optimal matching of the UAVs with the rescue leaders (Section 5.4.3). We consider $\tau_h = 0.985$ sec, $\tau_t = 0.015$ sec, $t = 1$ sec, $P_u^{max} = 85$W, $d_{v,v'} \in [30, 350]$m, $\lambda = 1$, $E_v \in [100, 400]$ J, $D_v \in [30, 350]$m, $W = 5 \cdot 10^6$Hz, $c = 1$, $b = 0.7$, $\mu = 0.5$, $FT_u \in (0, 1]$, and $R_u \in [30, 350]$m. We consider $|V| = 100$ victims, unless otherwise stated. The proposed framework's evaluation was conducted in a HP Laptop, 1.8GHz Intel Core i7, 16GB LPDDR3 RAM.

5.4.1 Impact of Socio-Physical Parameters

Three comparative scenarios regarding the victims' socio-physical characteristics are evaluated: (i) *Best*: victims with high communication interest reside close to each other; (ii) *Worst*: victims with high communication interest reside far away from each other; and (iii) *Random*: victims have random communication interest and

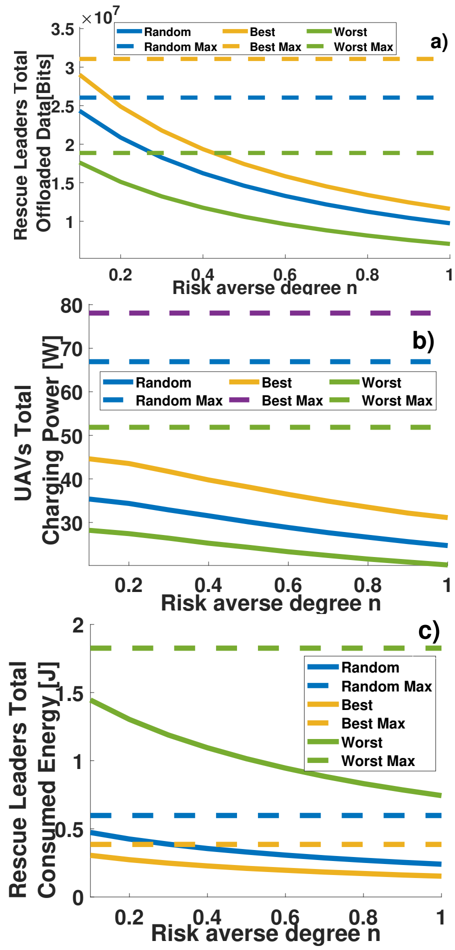

Figure 5.3: (a) Rescue leaders' total amount of offloaded data, (b) UAVs' total charging power, and (c) Rescue leaders' total consumed energy w.r.t. risk averse degree for various comparative scenario

distance among each other. Fig. 5.2a-5.2d present the victims' data offloaded to their rescue leaders, their corresponding total consumed energy, the number of created rescue groups, and their corresponding average size, respectively, as a function of the number of victims for the three considered comparative scenarios. The results reveal that under the best case scenario, few homogeneous (in terms of the victims' socio-physical characteristics) rescue groups are created (Fig. 5.2c) of large average size (Fig. 5.2d), while the victims achieve to offload a large amount of data (Fig. 5.2a) with small consumed energy (Fig. 5.2b), due to their close proximity and good channel conditions among each other. The exact opposite holds true in the worst case scenario, while the random scenario presents an intermediate behavior between the best and worst case scenarios.

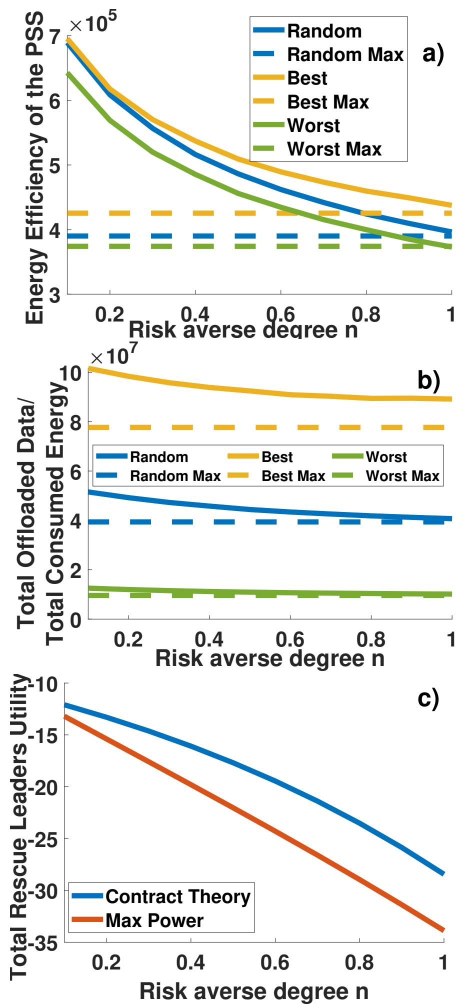

Figure 5.4: (a) Energy Efficiency of the PSS, (b) Ratio of the rescue leaders total offloaded data over the total consumed energy, and (c) Total rescue leaders utility w.r.t. risk averse degree for various comparative scenario

5.4.2 The Benefits of Contract Theory

We also present the impact of the victims' risk-aware behavior on the resource management and the benefits of adopting contract theory to model the interactions among the UAVs and the rescue leaders. Six comparative scenarios are considered; three based on the proposed contract-theoretic resource control approach while assuming the best, worst, and random scenarios (Section 5.4.1), and the corresponding three scenarios that conclude by assuming that the UAVs charge the rescue leaders' devices with their maximum available power (referred to as Best Max, Worst Max, and Random Max respectively). Fig. 5.3a-5.3c present the rescue leaders' total amount of offloaded data, the UAVs' total charging power, and the rescue leaders' corresponding consumed energy to offload their data to the UAVs, respectively, as a function of the rescue leaders' risk averse degree, for all the considered comparative scenarios. It is observed that, with reference to the contract-theoretic based scenarios, as the rescue leaders become more risk averse (i.e., high value of the risk averse degree n), they tend to invest less effort in terms of offloading their data to the UAVs (Fig. 5.3a), thus, they consume less energy in their data transmission (Fig. 5.3c) and enjoy less rewards (i.e., charging power) from the UAVs (Fig. 5.3b). Also, in the comparative scenario, where the UAVs provide their maximum available charging power (Fig. 5.3b) to incentivize the rescue leaders to offload more data (Fig. 5.3a), this goal is achieved by immensely sacrificing the energy efficiency of the public safety system (PSS), as shown in Fig. 5.4a.

Specifically, Fig. 5.4a depicts the PSS's energy efficiency defined as the total amount of offloaded data by the rescue leaders over the corresponding spent charging power by the UAVs as a function of the rescue leaders' risk averse degree n. The results reveal that the UAVs' charging power is not well-spent, when they charge the rescue leaders with their maximum available charging power, and the UAVs' energy cost for every unit of collected information is higher for any examined topology of the PSS and the victims' socio-physical characteristics. This, demonstrates the benefit of the contract-theoretic control of the resources from the PSS's point of view. Moreover, the proposed framework is also valuable for the rescue leaders, as it enables them to achieve greater utility (Eq. 5.1) compared to the scenario of having their devices charged with the UAVs' maximum charging power (Fig. 5.4c).

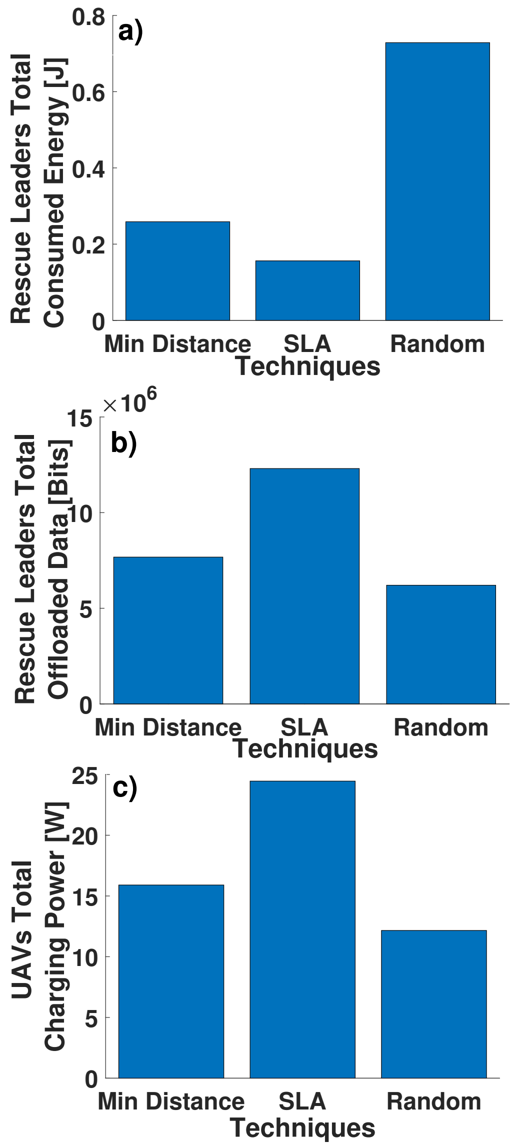

Figure 5.5: Reinforcement learning-based matching between the UAVs and the rescue leaders – A comparative evaluation

5.4.3 Intelligent Matching between UAVs & Rescue Leaders

In this subsection, we highlight the benefits of adopting a reinforcement learning mechanism to enable the rescue leaders to optimally select a UAV to offload their data, while considering the characteristics of the PSS. Two indicative alternative approaches are also considered for comparison purposes: (a) *Min Distance*: the rescue leaders offload their data to the closest UAV; and (b) *Random*: the rescue leaders randomly select a UAV to offload their data. Fig. 5.5a-5.5c illustrate the rescue leaders' total consumed energy, their corresponding total amount of offloaded data, and the UAVs' total charging power, respectively, for the considered comparative scenarios. The results reveal that the reinforcement learning approach enables the rescue leaders to thoroughly learn their surrounding environment and make a sophisticated choice of a UAV, as indicated by the holistic reputation function (Eq. 5.9). Thus, the rescue leaders achieve to report a larger amount of data (Fig. 5.5b), compared to the other comparative scenarios, while consuming the lowest amount of energy (Fig. 5.5a), and enjoying greater charging power from the UAVs (Fig. 5.5c).

5.5 Conclusions

In this chapter, a resource orchestration framework is introduced in a UAV-assisted WPCN within a public safety system, based on the principles of contract theory and reinforcement learning. The key objective and novelty of this framework is that it enables the energy efficient information acquisition from the victims, while considering their risk-aware behavior. Detailed numerical results, obtained through modeling and simulation, demonstrate the benefits and superiority of the proposed framework in terms of energy-efficiency, information acquisition from the disaster area, and intelligent users' incentivization to support the rescue operation. Our future research plans include the extension of the proposed framework to consider the backhaul connection between the UAVs and the emergency control center, thus offering an energy efficient end-to-end data acquisition and transmission solution.

Chapter 6

Health Data Acquisition from Wearable Devices during a Pandemic

6.1 Introduction

The last century has seen a plethora of pandemics, such as the Spanish flu (1918), the Hong Kong flu (1968) and the Swine flu (2009) [124]. Towards controlling the spread of a pandemic, the collection of citizens' health data, such as heart rate, body temperature, arterial oxygen saturation, from everyday wearable devices (e.g., smart phones, fitness trackers, etc.) is critical for a large number of applications. Some indicative applications include contact tracing, rapid diagnosis, patients' remote monitoring, and reducing the workload of the medical industry [125]. Furthermore, during a pandemic outbreak, except for the impact on humans' health, the global economy is also heavily impacted, while the non-essential services are forced to shut down. The recent pandemic of COVID-19 is expected to create a loss of 5.5 trillion US dollars in the next two years [126]. This chapter aims at jointly tackling the problem of citizens' health data acquisition from wearable devices and the economic survival of the businesses during a pandemic, by introducing a novel techno-economics approach.

6.1.1 Related Work

The collection of health data during pandemics, is crucial regarding both the healthcare planning, as well as towards strategizing for the sustainability of the economy [127]. Due to the recent COVID-19 outbreak, several wearables have been designed and exploited to gather citizens' health data. WHOOP Strap 3.0 is a wrist-mounted wearable measuring cardiorespiratory variations and reporting them via Bluetooth to the citizen's smart phone [128]. Estimote's wearable devices enable the wearer to update its health status and share it along with its location to a centralized entity, e.g., employer [129]. Biosensor Patch 1AX measures the body temperature, respiration rate, electrocardiogram trace, and heart rate and report them to the person's smartphone over Bluetooth [130]. All those sensors, along with other wearables, have been recently employed to deal with the COVID-19 outbreak. However, humans are hesitant in providing their health data, thus, sophisticatedly designed extrinsic and/or intrinsic incentive mechanisms should be designed [131].

In addition to the healthcare planning during pandemics, the economic impact and the change of the citizens' consumption culture play a key role on the sustainability of the economy [132]. McKinsey & Co. released a study showing that more than 70% of the US population has reduced its purchases in personal care services, hotel stays, out of home entertainment, restaurants, and others, during the COVID-19 era [133]. A pandemic has a devastating economic impact on the automotive, aviation, tourism, oil, construction, food, healthcare and medical industries [134]. Thus, creating the proper conditions to enable the citizens to keep their consumption culture, while respecting all the health protection rules, is important to move the economy during and past the pandemic.

6.1.2 Contributions & Outline

Despite the efforts made in previous works, in regards to the data acquisition from wearable devices during a pandemic and the economy's sustainability, the solution of those two problems still remains highly fragmented. Moreover, the incentivization of the citizens to provide their health data to the healthcare operator, e.g., operating at a district level, is even more challenging. In this work, we strive to tackle exactly these issues, by introducing a behavioral and labor economics based approach. The main contributions of this research work are summarized below.

1. A smart city scenario consisting of the healthcare operator, multiple businesses, and citizens with wearable devices is considered (Section 6.2). The citizens select to visit the businesses and perform purchases by exploiting the physical and social characteristics of both themselves and the businesses, based on a reinforcement learning approach (Section 6.3).

2. A behavioral economics approach is introduced to incentivize the citizens to provide the wearable devices' data to the businesses, while the latter provide extrinsic motivation to the citizens to facilitate their engagement in performing purchases. An optimization problem is formulated and solved to determine each citizen's optimal amount of reported data and the corresponding received rewards by the selected business (Section 6.4).

3. Given the overall data collected at each business, a labor economics-based approach is proposed to enable the healthcare operator at the district level to incentivize the businesses with monetary rewards to report the citizens' collected data to

it for further exploitation regarding the healthcare planning. The healthcare operator's optimal provided monetary rewards to the businesses, and the optimal amount of reported data by the latter ones are determined via formulating the problem as a contract-theoretic optimization problem (Section 6.5). Creating the aforementioned three layers approach, the joint goal of data acquisition and economy's sustainability is achieved.

4. A series of simulation experiments demonstrate the performance, effectiveness, and robustness of the overall proposed framework under various realistic scenarios (Section 6.6).

Finally, Section 6.7 concludes the chapter.

6.2 System Model

A smart city environment is considered consisting of the sets of citizens denoted as $C = \{1, \ldots, c, \ldots, |C|\}$, businesses (e.g., restaurants, hotels, bookstores) $S = \{1, \ldots, s, \ldots, |S|\}$, and the city's healthcare operator. The citizens equipped with wearable devices, can collect and provide personal health data. Those data are private, thus, the citizens should be incentivized to provide them, if they choose so. The healthcare operator is interested in collecting those data to facilitate the smart city's health protection planning during a pandemic. The businesses suffer from economic losses during such a pandemic, thus, they are interested in incentivizing the citizens to visit them, while respecting the announced health protection rules. The businesses act as a liaison between the citizens and the healthcare operator, in order to collect the citizens' health data from the wearable devices, while at the same time guaranteeing their own economic sustainability. In particular, the businesses offer personalized rewards to the citizens (e.g., coupons, discounts), and the citizens provide as an exchange their health data. Then, the healthcare operator provides personalized rewards (e.g., tax reduction) to the businesses, while the latter provide the collected health data from the citizens that visited them.

6.3 Reinforcement Learning-based Selection

A socio-physical based approach is introduced to enable the citizens to select the business that they will visit via exploiting the principles of reinforcement learning (RL). Each business is characterized by its reputation $R_s \in (0,1]$. Moreover, let us denote by $ph_s \in (0,1]$ the popularity $ph_s \in (0,1]$ of each business at an examined time slot (the latter can be easily retrieved from its Google reviews' profile, in terms of how busy is the business with customers). Each citizen c has a social interest $SI_{c,s} \in (0,1]$ to visit a business s. We denote by $|C|_s$ the number of citizens selecting to visit a business s at a time slot (e.g., one hour) that the system is examined.

A citizen c receives a reward (i.e., personal satisfaction) defined as: $r_{c,s} = \frac{SI_{c,s} R_s}{|C|_s ph_s}$ by visiting a business s, taking into account that a citizen wants to visit a reputable business of high personal interest, which however is not very crowded in order to easily respect the health protection rules. The corresponding normalized reward is $\hat{r}_{c,s} = r_{c,s} / \sum_{s=1}^{|S|} r_{c,s}$. The citizens can act as stochastic learning automata (SLA) making distributed decisions about themselves aiming to optimize their long-term reward by visiting a business. Based on the iterative SLA reinforcement learning algorithm, each citizen selects to visit the same (Eq. 6.1a) or a different business (Eq. 6.1b) based on the following probabilistic rule [135]:

$$P_{c,s}^{(ite+1)} = P_{c,s}^{(ite)} + b\hat{r}_{c,s}(1 - P_{c,s}^{(ite)}), s_c^{(ite+1)} = s_c^{(ite)} \tag{6.1a}$$

$$P_{c,s}^{(ite+1)} = P_{c,s}^{(ite)} - b\hat{r}_{c,s} P_{c,s}^{(ite)}, s_c^{(ite+1)} \neq s_c^{(ite)} \tag{6.1b}$$

where $0 < b < 1$ is the learning rate. For large values of b, the citizens do not thoroughly explore their available options, thus, the algorithm converges fast, while the opposite holds true for small values of b. The SLA algorithm converges when for each citizen there is a selection to visit a business with probability close to one ($P_{c,s}^{(ite)} \rightarrow 1$). The SLA algorithm can be implemented in a mobile application and help the citizens making personal decisions regarding the businesses that they will visit during a pandemic.

6.4 Contract-theoretic Data Collection

The problem of the citizens' incentivization by the businesses to provide their health data via offering personalized rewards is studied based on the principles of behavioral economics. Each business may provide M_s [\$] maximum rewards to a citizen, while $x_c \in [0, 1]$ denotes the actual portion of M_s to be offered to the citizen as an exchange to the actual reported health data. Each citizen can provide d_c [bits] total amount of data captured by its wearable devices, while $y_c \in [0, 1]$ denotes the percentage of them reported to the business. Naturally, each citizen has some personal behavioral characteristics in terms of how hesitant it is to report its personal health data. The citizen's risk aversion behavior to provide its data is captured through the parameter $n_c \in (0, 1]$. The greater the citizen's risk aversion n_c is, the more conservative is the citizen in reporting its data.

Following the principles of behavioral economics [39], the citizen's utility is represented, as follows:

$$U_c(x_c, y_c) = -e^{-n_c[x_c - \psi(y_c)]} \tag{6.2}$$

where $\psi(y_c)$ is the citizen's cost function to report its data to the business capturing for example its personal wearable devices' energy consumption, data usage, etc. The cost function is strictly increasing with respect to the percentage y_c of the reported data, e.g., $\psi(y_c) = \frac{cy_c^2}{2}$, where $c > 0$ is a constant cost factor.

It is noted that when a citizen provides $y_c d_c$ data, not all of its data are useful for the healthcare operator, thus, we consider the citizen's contribution in terms of data as $q_c = y_c + \epsilon$, where ϵ represents some noisy data. The variable ϵ follows a normal distribution with zero mean and variance σ^2. Also, each business is considered to provide a percentage reward $x_c = f_s + v_c q_c$, where f_s is a fixed reward and v_c is a variable compensation related to the citizen's provided data.

Towards determining the pair of optimal reward and optimal portion of reported data, i.e., $\{x_c^* M_s, y_c^* d_c\}$, the following optimization problem is formulated aiming at: (1) maximizing each business's expected benefit/profit (Eq. 6.3a), (ii) guaranteeing a minimum satisfaction $U_{c,min}$ for each citizen (Eq. 6.3b), thus capturing its individual rationality (IR), and (iii) maximizing the citizen's achieved satisfaction (Eq. 6.3c) to be incentive compatible (IC) to participate in the data reporting process.

$$\max_{\{x_c, y_c\}_{\forall c \in C}} \quad \mathbb{E}(q_c - x_c) \tag{6.3a}$$

$$\text{s.t.} \quad \mathbb{E}(-e^{-n_c[x_c - \psi(y_c)]}) \geq U_{c,min} \quad \text{(IR)} \tag{6.3b}$$

$$y_c \in \underset{y_c}{\operatorname{argmax}}\, \mathbb{E}(-e^{-n_c[x_c - \psi(y_c)]}) \quad \text{(IC)} \tag{6.3c}$$

The citizen's expected utility can be rewritten when given that $\mathbb{E}(e^{-n_c v_c \epsilon}) = e^{\frac{n_c v_c^2 \sigma^2}{2}}$ as $\mathbb{E}(-e^{-n_c[x_c - \psi(y_c)]}) = -e^{-n_c[f_s + v_c y_c - \frac{cy_c^2}{2} - \frac{n_c v_c^2 \sigma^2}{2}])}$. Thus, by solving (6.3c), we have:

$y_c^* d_c = v_c d_c / c$, and, we can rewrite the problem (6.3a)-(6.3c).

$$\max_{\{x_c, y_c\}_{\forall c \in C}} [\frac{v_c}{c} - (f_s + \frac{v_c^2}{c})] \tag{6.4a}$$

$$\text{s.t.} \quad f_s + \frac{v_c^2}{c} - \frac{v_c^2}{2c} - \frac{n_c v_c^2 \sigma^2}{2} = x_c \tag{6.4b}$$

The solution of the problem (6.4a)-(6.4b) can be easily determined as $x_c^* M_s = [f_s + \frac{1}{1+n_c c\sigma^2}(\frac{v_c}{c} + \epsilon)]M_s$. Thus, the contract of optimal reward and optimal portion of reported data, i.e., $\{x_c^* M_s, y_c^* d_c\}$, is determined.

6.5 Healthcare Operator's Data Acquisition

In this section, a contract-theoretic framework is introduced in order the healthcare operator to incentivize the businesses to provide the collected health data to it by providing tailored rewards to them. Each business has collected in a time slot (e.g., one hour) a total amount of data $\sum_{c=1}^{|C|_s} y_c^* d_c$ [bits]. Thus, we define the business's type as, $t_s = \frac{\sum_{c=1}^{|C|_s} y_c^* d_c}{\max_{\forall s \in S}\{\sum_{c=1}^{|C|_s} y_c^* d_c\}} \in [0, 1]$, reflecting the value of each business to the healthcare operator in terms of its ability to provide data. Each business performs an effort $q_s \in [0, 1]$ in terms of reporting the portion of data $q_s \sum_{c=1}^{|C|_s} y_c^* d_c$ to the healthcare operator, and receives a personalized reward $r_s(q_s) = t_s q_s$ from the operator, accounting for both the business's amount of reported data, as well as its potential to report a large amount of data captured by its type.

Each business's utility presents its profit/benefit from reporting the collected data and is given as $U_s(q_s) = t_s e(r_s) - q_s$, where $e(r_s)$ is the evaluation function. The latter function is a strictly increasing and concave function with respect to the business's received reward, e.g., $e(r_s) = \sqrt{r_s}$, and captures the way the business evaluates the received monetary rewards. The healthcare operator is unaware of the amount of data that each business may have collected, thus, it estimates the type of each business with probability Pr_s, which in this chapter, without loss of generality, is assumed to follow a uniform distribution. The profit/benefit of the healthcare operator from collecting the data, while providing tailored rewards to the businesses is given as $U_o(\mathbf{q}) = \sum_{s=1}^{|S|} Pr_s(q_s - \lambda r_s)$, where $\mathbf{q} = \{q_1, \dots, q_{|S|}\}$. The social welfare of the overall system is $SW(\mathbf{q}) = U_o(\mathbf{q}) + \sum_{s=1}^{|S|} U_s(q_s)$.

Following the principles of Contract Theory, the healthcare operator aims at maximizing its benefit from the data acquisition process (Eq. 6.5a), while guaranteeing that the optimal contracts, i.e., $\{r_s^*, q_s^*\}$, are acceptable by the businesses. Towards achieving the latter goal, each business's profit should be non-negative to satisfy its personal individual rationality (Eq. 6.5b), the reward provided to each business should be tailored to its type in order to be incentive compatible for the business (Eq. 6.5c), and fairness should be guaranteed in the overall rewarding process. The latter means that a business of higher type, i.e., $t_1 < \dots < t_{|S|}$, will provide more data, i.e., $q_1 < \dots < q_{|S|}$, thus, it will receive a greater reward (Eq. 6.5d). The businesses are sorted with respect to their type for simplicity in the presentation. The corresponding contract-theoretic optimization problem is defined as follows.

$$\max_{\{r_s, q_s\}_{\forall s \in S}} U_o(\mathbf{q}) = \sum_{s=1}^{|S|} Pr_s(q_s - \lambda r_s) \tag{6.5a}$$

$$\text{s.t.} \quad t_s e(r_s) - q_s \geq 0, \forall s \in S \quad \text{(IR)} \tag{6.5b}$$

$$t_s e(r_s) - q_s \geq t_s e(r_s') - q_s', s \neq s', \forall s, s' \in S \quad \text{(IC)} \tag{6.5c}$$

$$0 \leq r_1 < \dots < r_s < \dots < r_{|S|} \tag{6.5d}$$

The above optimization problem is non-convex, thus, towards solving it, we will reduce its constraints. Initially, focusing on the IR constraints in Eq. 6.5b, and based on the IC constraint we have that $t_s e(r_s) - q_s \geq t_s e(r_s') - q_s' \geq t_s e(r_1) - q_1$, while the last inequality holds true given that the evaluation function is strictly increasing with respect to r_s, and $r_1 < \cdots < r_{|S|}$ and $q_1 < \cdots < q_{|S|}$. Also, given that $t_1 < \cdots < t_{|S|}$, we have $t_s e(r_1) - q_1 \geq t_1 e(r_1) - q_1 \geq 0$. Moreover, the healthcare operator will provide the sufficient reward to just incentivize the businesses to participate in the data acquisition process, thus, Eq. 6.5b can be equivalently substituted with $t_1 e(r_1) - q_1 = 0$.

Towards reducing the IC constraints presented in Eq. 6.5c, we adopt the following terminology: (i) $s, s', s' \in \{1, \ldots, s-1\}$ downward IC (DIC), (ii) $s, s-1, \forall s, s-1 \in S$ local downward IC (LDIC), (iii) $s, s', s' \in \{s+1, \ldots, |S|\}$ upward IC (UIC), and (iv) $s, s+1, \forall s, s+1 \in S$ local upward IC (LUIC) constraints [39].

Proposition 1. *All the DIC constraints can be captured by the LDIC constraints.*

Proposition 2. *All the UIC constraints can be captured by the LDIC constraints.*

Based on the above analysis of the reduction of the IR and IC constraints, we can rewrite the optimization problem present it in (6.5a)-(6.5d), as follows.

$$\max_{\{r_s, q_s\}_{\forall s \in S}} U_o(\mathbf{q}) = \sum_{s=1}^{|S|} Pr_s(q_s - \lambda r_s) \tag{6.6a}$$

$$\text{s.t.} \quad t_1 e(r_1) - q_1 = 0, \forall s \in S \quad \text{(IR)} \tag{6.6b}$$

$$t_s e(r_s) - q_s = t_s e(r_{s-1}) - q_{s-1}, \forall s, s-1 \in S \quad \text{(IC)} \tag{6.6c}$$

$$0 \leq r_1 < \cdots < r_s < \cdots < r_{|S|} \tag{6.6d}$$

The resulting optimization problem in (6.6a)-(6.6d) can be easily solved using standard tools of convex optimization and determine the optimal contracts $\{r_s^*, q_s^*\}$. Consequently, the total amount of data collected by the healthcare operator is: $\sum_{s=1}^{|S|} q_s^* \sum_{c=1}^{|C|} y_c^* d_c$.

6.6 Numerical Results

In this section, we evaluate the performance of the proposed economic-driven health data acquisition framework via numerical simulations and detailed comparative analysis. Initially, the impact of the socio-physical parameters in the citizens' business selection and data acquisition is presented under different comparative scenarios (Section 6.6.1). The risk-aware citizens' behavior is analyzed under different cost parameters to show their correlation in the data collection process and the critical role of the citizens involvement during a pandemic (Section 6.6.2). Finally, a detailed comparative evaluation is performed to illustrate the businesses' and healthcare operator's interactions to jointly achieve the economic survival of the first and the data collection of the latter towards performing the healthcare planning in the smart city (Section 6.6.3). In the rest of the analysis, we consider $|C| = 500$ citizens, $M_s \in [2, 4]\$$, $d_c \in [2, 3.5]$ Mbits, $|S| = 15$ businesses, $R_s \in [0, 1]$, and $SI_{c,s} \in [0, 1]$ and $ph_s \in [0, 1]$ with increasing values with respect to the business's ID and c = 2, unless otherwise explicitly stated. The proposed framework's evaluation was conducted in a HP Laptop, 1.8GHz Intel Core i7, 16GB LPDDR3 RAM.

6.6.1 Socio-physical-based Business Selection

In this subsection, the performance of the reinforcement learning-based business selection by the citizens in an autonomous manner via considering their social-physical characteristics (Section 6.3) is studied. In particular, the proposed SLA distributed decision-making under various learning rate values, i.e, $b = \{0.1, 0.4, 0.7, 0.9\}$, is compared against (a) *Random Scenario*: the citizens select randomly which business to visit; and (b) *Congestion Scenario:* the citizens act as SLA with reward function $r_{c,s} = 1/|C|_s$ aiming to visit the less congested business without exploiting their social-physical preferences. For fairness in the comparison, for all examined alterna-

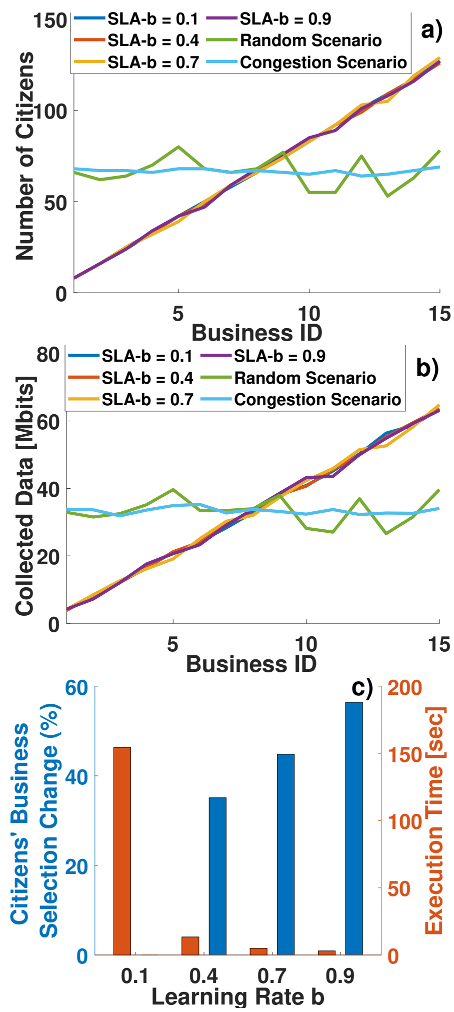

Figure 6.1: Reinforcement Learning-based Business Selection – A Comparative Evaluation

tives, we apply the behavioral contract-theoretic data collection approach (Section 6.4). Fig. 6.1a and Fig. 6.1b present the number of citizens that visited each business and the amount of data collected by the latter. The results reveal that under the SLA approach, the citizens successfully select the business that provides them higher reward ($r_{c,s} = \frac{SI_{c,s}R_s}{|C|_s ph_s}$) given the superior social-physical characteristics. Also, the corresponding businesses that better satisfy the citizens' social-physical needs, achieve to collect more data that can be exploited for making further profit by reporting them to the healthcare operator. Thus, the businesses that offer better rewards

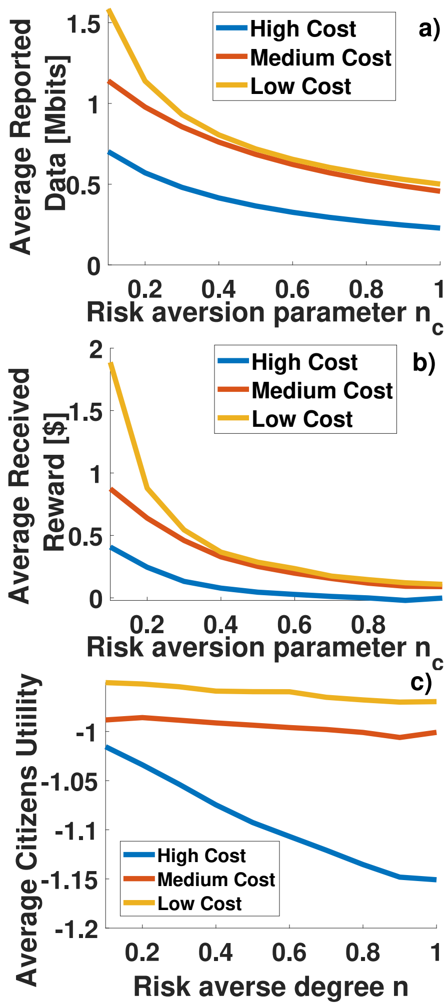

Figure 6.2: Behavioral and Cost-based Data Collection

$r_{c,s}$ to the citizens have dual economic benefit, as they attract more customers and collect more data from them that can exchange with the healthcare operator for additional monetary rewards. On the other hand, the random selection allocates the citizens to the businesses in a non-sophisticated manner. The latter has a result the heterogeneous congestion of the businesses (Fig. 6.1a) and the unbalanced data collection (Fig. 6.1b). Also, the Congestion Scenario tends to blindly balance the number of citizens allocated to the businesses, without accounting for any social parameters, thus, resulting in lower rewards $r_{c,s} = \frac{SI_{c,s}R_s}{|C|_s ph_s}$ compared to the SLA approach.

Focusing on the pure operation of the SLA approach, Fig. 6.1c presents the percentage of changes of the citizens' selections to visit a business compared to their selection for learning rate $b = 0.1$. The latter is considered as the ground-truth given that the citizens exploit thoroughly their choices for small values of the learning rate concluding to higher values of reward by their decision to visit a business. Also, Fig. 6.1c shows the real execution time of the SLA approach until it converges to a stable decision. The results show that as the value of the learning parameter increases, the citizens make faster decisions, however, they deviate more from the ground-truth, consequently receiving smaller rewards.

6.6.2 Behavioral and Cost-based Data Collection

Subsequently, we turn our attention to the quantification of the impact of the citizens' risk-aware behavior and their experienced cost to report their data to the business that they select to visit, on their final decisions regarding the amount of data that they report.

Specifically, Fig. 6.2a-6.2c present the average values of the citizen's reported health data, their received reward from the businesses, and their achieved utility (Eq. 6.2) as a function of the citizen's risk aversion parameter n_c. Three comparative scenarios are considered regarding the cost that the citizens experience to report their data: (a) *High Cost*: $\psi(y_c) = \frac{cy_c^2}{2}, c = 3$, (b) *Medium Cost*: $\psi(y_c) = \frac{cy_c^2}{2}, c = 2$, and (c) *Low Cost*: $\psi(y_c) = \frac{cy_c^4}{4}, c = 2$, where $y_c \in [0, 1]$. The results reveal that as the citizens become more risk averse (i.e., increasing value of the risk aversion parameter n_c) in terms of reporting their health data, they tend to actually report a smaller amount of data (Fig. 6.2a), thus, they receive a smaller reward from the businesses (Fig. 6.2b), resulting finally in a decreased enjoyed utility (Fig. 6.2c). In other words, the more hesitant the citizens become with sharing their wearable devices'

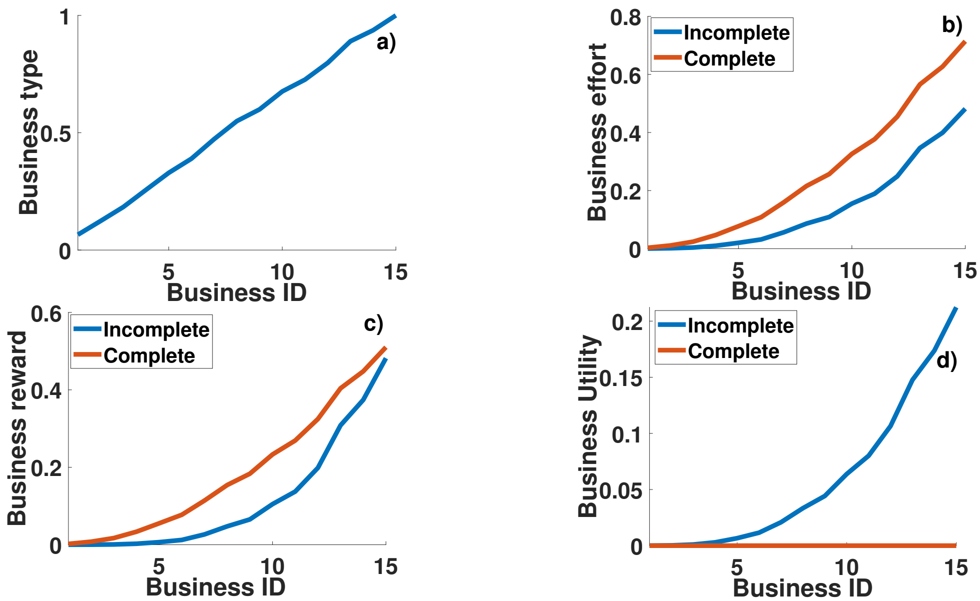

Figure 6.3: Impact of socio-physical parameters under different comparative scenarios

health data, the less rewards they experience from the businesses.

Focusing on the impact of the cost that the citizens experience to report their data to the businesses, as expected the results show that the higher the cost, the less data the citizens report to the businesses that they visit (Fig. 6.2a). Thus, the businesses offer lower rewards to the citizens (Fig. 6.2b), and the latter achieve less satisfaction (Fig. 6.2c). Based on the overall above behavioral and cost analysis, we conclude that a citizen who is less hesitant to report its data, and experiences low cost, has the potential to contribute more in the data acquisition process and the businesses' economic sustainability.

6.6.3 Businesses and Healthcare Operator Interactions

We further explore the interactions among the businesses and the healthcare operator under the case of complete and incomplete information availability, i.e., the healthcare operator knows the businesses types in a deterministic and probabilistic manner,

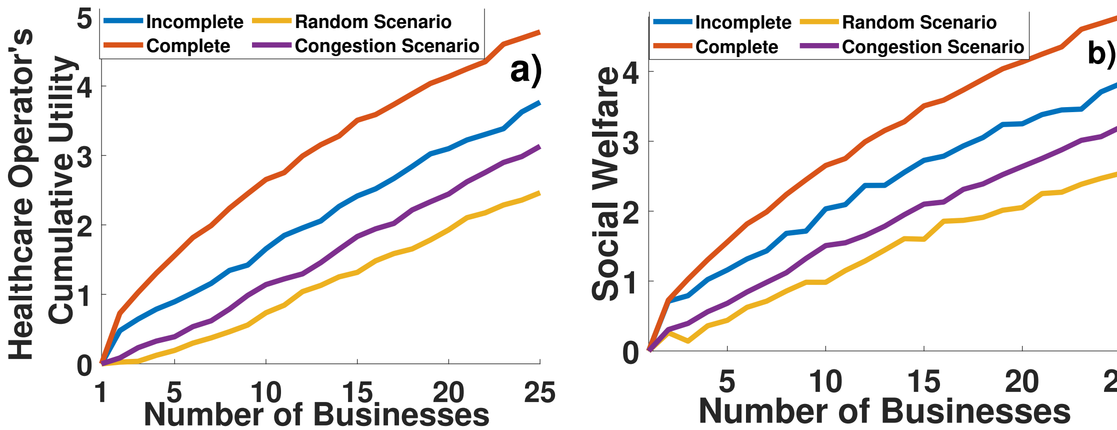

Figure 6.4: System's Operation - A Comparative Analysis

respectively. Moreover, a comparative evaluation of our proposed contract-theoretic framework against other approaches is performed, to reveal our framework's benefits.

In particular, Fig. 6.3a presents the businesses' type t_s as a function of their ID. Fig. 6.3b-6.3d illustrate the businesses' normalized effort q_s, their normalized received reward r_s from the healthcare operator and their achieved utility U_s as a function of the business ID, respectively, for the complete and incomplete information scenarios. The results reveal that as the business's type increases (i.e., the business has collected more data from the citizens), the greater effort it invests to report the data to the healthcare operator (Fig. 6.3b). This results in the business receiving a larger reward (Fig. 6.3c) and consequently a higher utility (Fig. 6.3d). Moreover, it is shown that higher reward and corresponding higher effort is observed under the complete information scenario, as the healthcare operator knows the businesses' types, thus, it provides just sufficient rewards to collect all their available data. For that reason, the businesses' utility is zero, as the healthcare operator offers that level of rewards to the businesses, in order the latter to just balance the cost to report their data with the received monetary rewards from the healthcare operator.

Subsequently, in Fig. 6.4a and Fig. 6.4b, a comparative evaluation of the healthcare operator's cumulative utility and the overall social welfare are presented, respectively, under the complete and incomplete information scenarios, and the Random and Congestion Scenarios (described in Section 6.6.1). The results reveal that the complete information scenario achieves the best results as the healthcare operator has full knowledge about the businesses' types and can more accurately plan the data acquisition process (Fig. 6.4a), while guaranteeing the high levels of satisfaction of all the involved entities (Fig. 6.4b). The incomplete information scenario, which corresponds to a more realistic implementation, achieves the second best results, achieving approximately 22% worse social welfare compared to the (ideal) complete information scenario. This observation indicates that the proposed framework behaves very well under the challenging scenario, where the healthcare operator is unaware of the businesses' characteristics. Finally, the non-sophisticated random scenario results in the worst results, while the congestion scenario results in unsatisfied citizens visiting businesses that they have low social interest to visit. Accordingly, in this case the citizens are not motivated to report a large amount of their data and the system performs poorly.

6.7 Conclusions

In this chapter, a techno-economics based approach is introduced to deal with the joint problem of health data acquisition from the citizens and economic sustainability of the businesses in a smart city. A reinforcement learning approach is proposed to enable the citizens to choose the most preferable business to visit by exploiting their socio-physical characteristics. A novel incentivization model is introduced in order for the citizens to provide their health data to the businesses that they visit, based on the behavioral economics, while improving the businesses' economic sustainability. Moreover, a contract-theoretic data acquisition framework is proposed enabling the healthcare operator to acquire the citizens' health data, towards facilitating the smart city's healthcare planning. Part of our current and future work involves the extension and realization of this model, by implementing the proposed frameworks in mobile Android and IOS applications, and test them in realistic environments under different scenarios of incomplete information.

Chapter 7

Museum and Visitors Interactions Enabled by Labor Economics

7.1 Introduction

Over the last decade, the world has witnessed a rapid growth of services and applications provided to citizens in order to improve and facilitate their everyday life activities. Indicative application domains include smart homes, transportation systems, education, smart cities, and cultural spaces, to name a few. To improve the Quality of Service (QoS) provided to the end-users, humans have become an integral part of the overall system design and the human factors are naturally taken into account during the system operation. Following this human-in-the-loop (HITL) approach, cyber-physical-social systems (CPSSs) have emerged combining the aspects of communications, computing, control, and human resources and attributes [136]. In this chapter, we focus our study on the cyber-physical-social system of a museum, instrumenting the interactions among the museum operator and the visitors, and incentivizing the latter to provide feedback to the former via a labor economics-based approach.

7.1.1 Related Work

In CPSSs, the participatory sensing, i.e., data collection from humans in order to analyze them and extract their characteristics and behaviors, is critical in designing and optimizing the CPSS to meet the humans' QoS and Quality of Experience (QoE) requirements. Various incentives have been introduced in recent literature in order to incentivize humans' participatory sensing [137]. Such incentives can be organized into two main classes: (a) monetary and (b) non-monetary incentives, enabling extrinsic and intrinsic human motivation, respectively [138]. The monetary incentives can be static, i.e., predetermined based on some criteria and unaffected by the changes in the CPSS, or dynamic, i.e., dependent on the real-time conditions of the CPSS. The non-monetary incentives can be: (i) collective, i.e., the humans are encouraged to work together for a common good; (ii) social, i.e., the humans are encouraged to achieve a social status; and (iii) socially interactive, i.e., the CPSS provides to the humans a feeling of social presence and belonging, such as in social networks and blogs [139].

In [140], a behavioral model is presented in order to categorize the humans into three types with respect to the participatory sensing process: malicious, speculative, and honest. An incentive mechanism is also introduced, which provides benefits to the humans in proportion to their reputation with respect to the collected information. In [141], a contract-theoretic participatory sensing mechanism is proposed. A participatory sensing platform announces various data acquisition tasks and provides benefits to humans based on their invested effort to collect and report the data. The ultimate goal of the participatory sensing in CPSSs is to improve their operation as dynamic and complex systems and, in parallel, to enhance the experienced QoE by the humans which play an active role therein [142].

Considering the cultural heritage space use case of CPSSs, such as museums, limited quantitative models have been introduced in the literature to jointly optimize the museum's operation and the visitors' QoE, while exploiting their interactions. In [143], the visitor's experience in a museum is studied via analyzing the information shared by the visitor to the social networks following the participatory sensing paradigm. The authors' goal is to exploit the collected information in order to identify the visitors' interests, and thus improve the exhibition's planning. The importance of participatory sensing and the corresponding generated big data has been highlighted in [144]. Specifically, an exploratory study was conducted with a participating group of museum professionals who identified the critical role of participatory sensing and big data in improvements to museum operations.

Furthermore, focusing on the visitors' QoE improvement, an empirical study was performed in [145] to identify the most influential parameters of visitors' QoE, while considering the visitors' style and characteristics. This work has been extended in [26], where a routing mechanism is introduced in order to recommend a museum tour to the visitors for the purpose of optimizing their QoE. In [146], the authors have implemented a light detection and ranging (LIDAR) system to detect and track the visitors' positions and mobility patterns in the museum. Based on the latter collected information, the visitors' behavior during the museum touring is inferred and used in designing the museum exhibition to improve visitor QoE. This research has been extended in [147], where the authors have performed a real experiment in the art gallery of the Ohara Museum of Arts in Japan to extract the visitors' behavioral characteristics during the museum visit.

Focusing on exploiting the visitors' behavioral patterns in order to improve their achieved QoE, Prospect Theory has been adopted in [148] to determine the optimal time that each visitor should spend per exhibit in order to optimize their QoE. This research work has been extended in [149] to accommodate the museum operator's goal of mitigating visitor congestion in the museum, while accounting for the visitors' goal to optimize their QoE. In the proposed model, sophisticated pricing mechanisms have been designed to enable the museum operator to incentivize the visitors to invest their visiting time appropriately, thereby jointly optimizing both the museum operator's and the visitors' aforementioned objectives. In [150], the authors focus on designing a mechanism of smart routing and recommendation provision for the visitors, while accounting for their behavioral and visiting styles. Also, in [151], a reinforcement learning approach is introduced in order to enable the visitors to perform a recommendation selection regarding their visiting style (e.g., map, facilitator, audiovisual equipment) in an autonomous manner. Moreover, a game-theoretic framework is proposed to determine their optimal visiting time in a distributed manner, so as to maximize their perceived QoE.

7.1.2 Contributions & Outline

In a nutshell, the existing research has focused on participatory sensing to infer the visitors' behavior in order to either optimize their QoE or enable the museum operator to perform optimal museum planning, such as with strategic exhibit deployment and congestion mitigation. However, very limited attention has been given to the joint accommodation of both the museum operator's and the visitors' goals, which still remains an open issue of high research and practical importance. In particular, the problem of orchestrating the museum operator's and the visitors' interactions while accounting for the visitors' behavioral characteristics, in an effort to jointly optimize the benefits of two entities with different and possibly diverse interests, is even more challenging. In this research work, we particularly strive to tackle these issues. The main contributions of this study that differentiate it from other literature are presented below.

1. The characteristics of the museum considered as a cyber-physical-social system

are identified. The visitors' characteristics and their visiting styles are studied in order to define the visitors' types. Each visitor is characterized by a unique type stemming from their level of knowledge with respect to the exhibition, their mobility pattern, their visiting style, and the time that they are willing to invest in the museum touring (Section 7.2).

2. A labor economics-based framework is introduced to capture the museum operator's and the visitors' benefits from the collected information, via the participatory sensing and the monetary incentives, respectively. Specifically, the museum operator provides monetary rewards to the visitors so that they will provide evaluations about the exhibits, while simultaneously accounting for the sensible spending of their time in the museum. The interactions among the museum operator and visitors are captured in appropriately designed utility functions following the principles of labor economics (Section 7.3).

3. A labor economics-based optimization problem is formulated to jointly consider and treat the museum operator's and the visitors' utilities. The problem is studied and solved under the scenarios of complete information (Section 7.4), wherein the museum operator knows the visitors' types deterministically, and incomplete information (Section 7.5), wherein the museum operator estimates the visitors' types probabilistically. The outcome jointly determines the visitors' optimal contributions, expressed in terms of their total provided evaluations of visited exhibits over their touring time, as well as the museum operator's optimal amount of rewards provided to each visitor.

4. With a detailed set of experiments, we show that the proposed approach based on principles of labor economics, outperforms a type-agnostic scenario which is unaware of the visitors' unique characteristics, with a 25% improvement upon the visitors' QoE. Our findings also show that the labor economics-based incentivization framework can provide a five-fold improvement of the social

welfare of the museum CPSS, when compared to the scenario in which the visitors provide the maximum possible number of evaluations and invest the minimum acceptable time for their tour (Section 7.6).

7.2 Museum: A Cyber-physical-social System

The cyber-physical-social system of a museum is considered consisting of the museum operator, who is responsible for the museum planning and management, and the set of visitors $N = \{1, \dots, n, \dots, |N|\}$. Each visitor has a maximum available time t_n^{Max} [min] that they are willing to invest in touring the museum. Also, the museum operator aims to collect feedback and evaluations from the visitors regarding their interest in the exhibits included in the exhibition. The number of evaluations provided by the visitor n is denoted as $E_n, E_n \in \mathbb{N}$. Furthermore, each visitor is characterized by their level of knowledge $k_n \in [0, 1]$ regarding the content of the exhibition. Based on their level of knowledge, the visitor's provided exhibit evaluations weigh accordingly.

Based on the seminal research work of Véron and Levasseur [152], the visitors are categorized in the visiting styles of ants, butterflies, grasshoppers, and fish, on the basis of their mobility pattern in the museum and time spent per exhibit. The ants visit all of the exhibits sequentially, spending similar time at each one. The butterflies visit almost all of the exhibits and spend varying times at each one. The grasshoppers spend a long time at each of a select few exhibits. Finally, the fish stand in the center of the room observing the majority of the exhibits without having any specific interest. Based on that animal metaphor of the visitor styles, we define the visitor's persona $p_n, p_n \in [0, 1]$, which expresses how trusted each visitor's evaluations are. The ant-persona more closely and carefully observes all of the exhibits, as compared to the butterfly-persona. The butterfly-persona is not biased to visit specific exhibits, as compared to the grasshopper-persona. The fish-persona can be thought of as the less-interested visitor at the exhibition, who may visit the museum chiefly to accompany their friends or family. Thus, the visitors' personas are ranked as $p_{fish} < p_{grasshopper} < p_{butterfly} < p_{ant}$, capturing the quality of evaluation that each visitor can provide.

Considering the visitors' characteristics, we define the visitor's type $\tau_n = \frac{p_n k_n}{t_n^{Max}} \in [0, 1]$, which jointly reflects the visitors' personas, their levels of knowledge, and their willingness to invest time in touring the museum. It is noted that the maximum visiting time t_n^{Max} is measured in minutes, and for all practical scenarios we consider $t_n^{Max} \geq 1min$.

7.3 Utility Functions: A Labor Economics Modeling

The principles of labor economics are adopted in order to model interactions among the museum operator and the visitors, as well as their respective benefits from museum planning and touring. Following the philosophy of labor economics, an "employer" provides personalized rewards to the "employees" in order to incentivize them to perform an action for the common good. The provided rewards are personalized based on the employee's contribution to the overall system's wellness. Usually, the employer is unaware of the employees' types (i.e., incomplete information scenario), which reflect their capability to provide contribution to the system. Thus, the rewards are provided to the employees in a probabilistic manner. On the other hand, the ideal scenario in which the employer knows the employees' types deterministically (e.g., from historical data) is used for benchmarking purposes. The latter scenario acts as the ground truth to evaluate the success of the devised model under incomplete information regarding the employees' types. The theory of labor economics aims to jointly optimize the employer's profit and benefits, while still satisfying the employees' QoE prerequisites, thereby concluding to a stable and rewarding equilibrium for the overall system.

Focusing on the museum CPSS use case scenario, the museum operator and the visitors act as the employer and the employees, respectively. The goal of the museum operator is to collect meaningful evaluations from the visitors to efficiently plan the exhibition, as well as to motivate the visitors to spend their time wisely during their tour. This goal is achieved by adopting an extrinsic motivation strategy and providing personalized monetary rewards to the visitors, such as ticket discounts and coupons for the museum's store. On the other hand, the visitors aim to satisfy their QoE prerequisites by enjoying the provided monetary rewards and investing their time meaningfully during their tour.

To begin capturing the above analysis, the visitor's contribution in the museum CPSS operation is defined as $x_n = \frac{E_n}{t_n}$ and, for clarity in the analysis and without loss of generality, is mapped to the interval $\hat{x}_n \in [0, 1]$. The museum CPSS benefits from an increased number of provided evaluations, E_n, in a considerate time, t_n [min], spent by the visitor for their tour. The museum operator provides personalized rewards, $r_n(\hat{x}_n) = \tau_n \hat{x}_n, r_n \in [0, 1]$, to the visitors which are dependent on their provided contribution, $\hat{x}_n$, and the quality of the information provided by their evaluation, as captured by their type, τ_n. Each visitor evaluates their personalized reward in terms of improving their perceived QoE via the evaluation function $e(r_n)$. The evaluation function is strictly increasing and concave with respect to the received rewards, e.g., $e(r_n) = \sqrt{r_n}$. The diminishing rate of return in reward perception is due to the fact that, especially after some point, visitors are constrained by their physical capability to further decrease their tour time and/or increase their number of provided evaluations.

Each visitor aims to improve their QoE, which consists of the visitor's evaluation of their personalized rewards based on their particular visiting type (first term of Eq. 7.1), while also considering the cost of providing their personal contribution to the smooth operation of the museum CPSS (second term of Eq. 7.1). Thus, the visitor's utility is formulated as follows:

$$U_n(\hat{x}_n) = \tau_n e(r_n) - k\hat{x}_n \tag{7.1}$$

where $k \in \mathbb{R}^+$ is the visitor's personal cost to provide their contribution, e.g., their smartphone's battery level expended to provide evaluations via the museum's mobile application.

On the other hand, as mentioned before, the museum operator benefits from collecting the visitors' contributions. Nevertheless, the museum operator is burdened with the cost of providing monetary rewards to the visitors in this attempt to extrinsically motivate them. Also, the visiting types $\tau_n, \forall n \in N$ of the visitors are unknown by the museum operator in the general case, so the latter one estimates the visitors' types with probability P_n, where $\sum_{n=1}^{|N|} P_n = 1$. Thus, the museum operator's utility is formulated as:

$$U_M(\hat{\mathbf{x}}) = \sum_{n=1}^{|N|} [P_n(\hat{x}_n - cr_n)] \tag{7.2}$$

where $c \in \mathbb{R}^+$ is the museum operator's cost to provide the rewards and $\hat{\mathbf{x}} = (\hat{x}_1, \dots, \hat{x}_n, \dots, \hat{x}_{|N|})$ is the visitor contribution vector.

Combining Eq. 7.1 and Eq. 7.2, the social welfare of the overall museum CPSS, consisting of the museum operator and the visitors, is given collectively as:

$$SW(\hat{\mathbf{x}}) = U_M(\hat{\mathbf{x}}) + \sum_{n=1}^{|N|} U_n(\hat{x}_n) \tag{7.3}$$

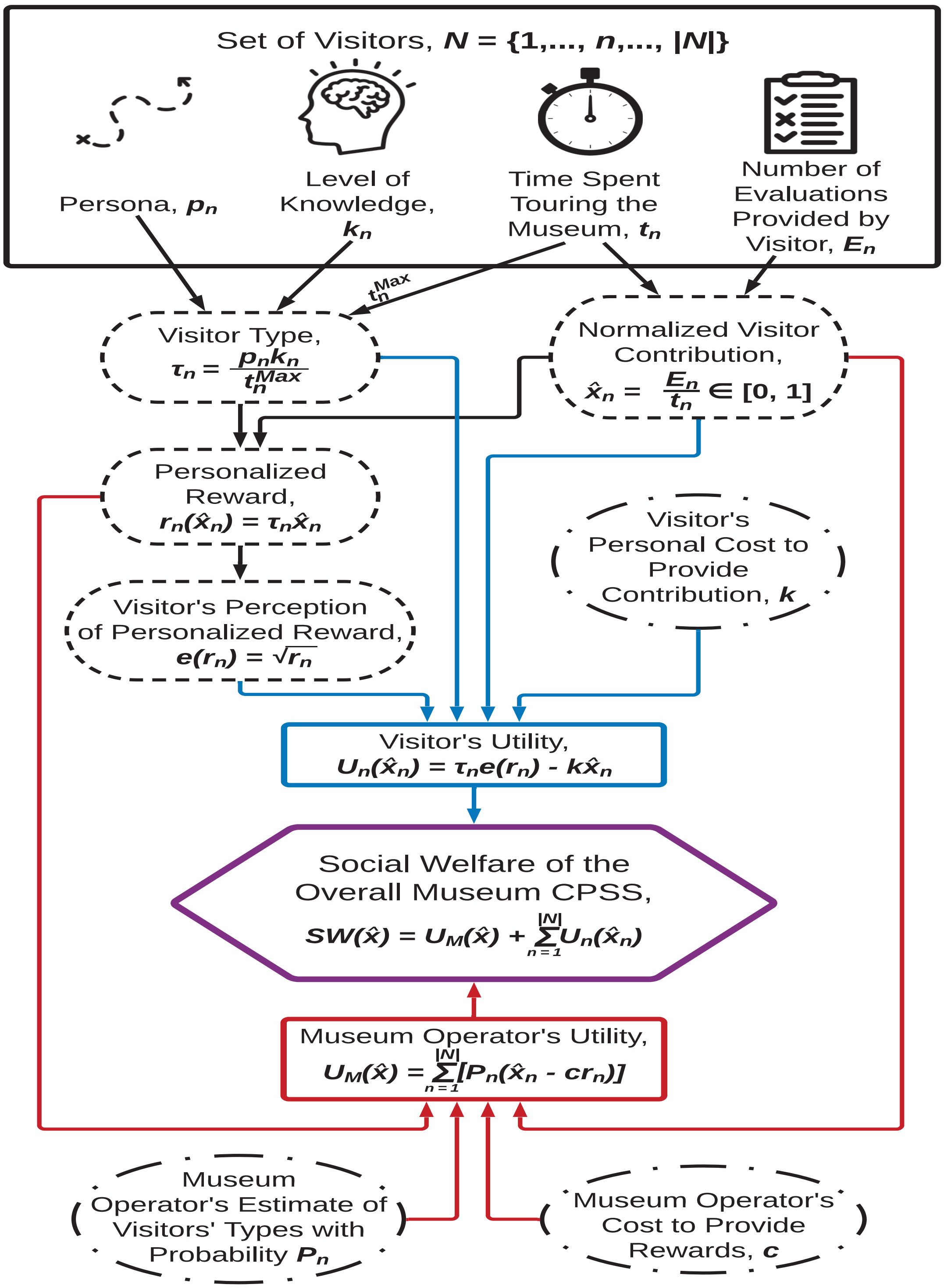

Figure 7.1: Museum and visitors interaction and feedback orchestration model.

A graphical representation of the analysis provided in Sections 7.2 and 7.3, reflecting the museum and visitors interaction and feedback orchestration model, is provided in Fig. 7.1.

7.4 Optimal Contracts under Complete Information

In this section, we examine the museum operator's and the visitors' interactions under the ideal case, with the museum operator knowing the visitors' types deterministically (i.e., complete information). This scenario is mainly used for benchmarking purposes. The goal of the museum operator is to optimize its profit and benefits, while jointly satisfying the visitors' QoE prerequisites, as they are captured by and reflected via their utility functions. This optimization problem, given the complete information of the visitors' types, can be formulated as:

$$\max_{\{r_n,\hat{x}_n\}_{\forall n\in N}} [U_M^n(\hat{x}_n) = \hat{x}_n - cr_n], \forall n \in N \tag{7.4a}$$

$$\text{s.t.} \quad \tau_n e(r_n) - k\hat{x}_n \geq 0 \quad \text{(IR)} \tag{7.4b}$$

where $U_M^n(\hat{x}_n)$ is the museum operator's utility due to each visitor of known type τ_n. The condition (7.4b) expresses the visitor's individual rationality (IR), which dictates that the visitor should experience at least a positive utility in order to be incentivized to interact with the museum operator.

The solution of the optimization problem (7.4a)-(7.4b) concludes to the optimal reward r_n^* and optimal contribution $\hat{x}_n^*$ for each visitor. The pair $\{r_n^*, \hat{x}_n^*\}$ is referred to as optimal "contract" in the remainder of our analysis.

Theorem 1. *(**Optimal Contract under Complete Information**): Under the complete information scenario, the optimal contract among the museum operator and each visitor $n, n \in N$ is $\{r_n^*, \hat{x}_n^*\} = \{(\frac{\tau_n}{2ck})^2, \frac{\tau_n{}^2}{2ck^2}\}$.*

Proof. The IR constraint in Eq. 7.4b can be reduced to $r_n = (\frac{k\hat{x}_n}{\tau_n})^2$, given that the museum operator will provide just-sufficient rewards to incentivize the visitors to participate in the smooth operation of the museum CPSS. Thus, by substituting the latter expression in Eq. 7.4a, taking the first order derivative with respect to $\hat{x}_n$, setting it to equal to zero, and solving with respect to $\hat{x}_n$, we have $\hat{x}_n^* = \frac{\tau_n{}^2}{2ck^2}$. Thus, we can easily derive that $r_n^* = (\frac{\tau_n}{2ck})^2$. □

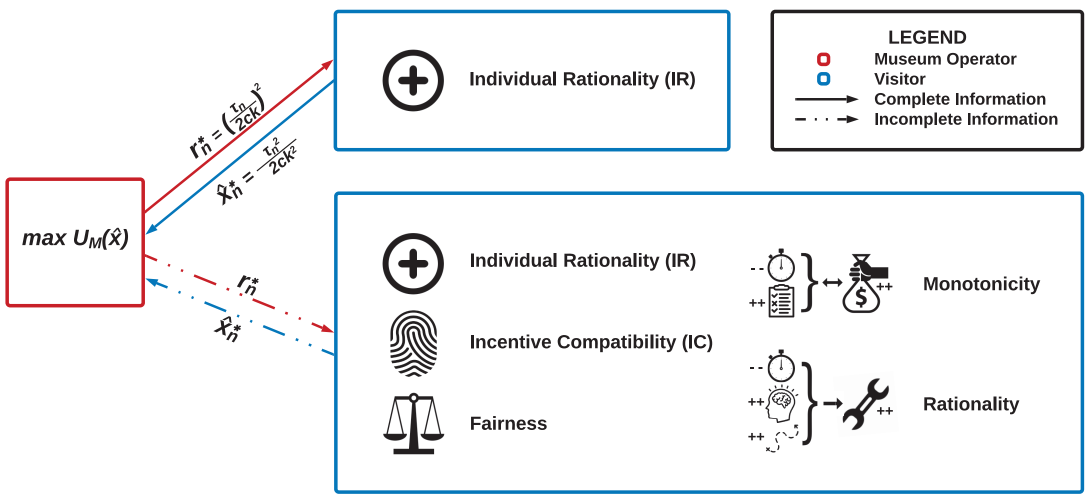

Figure 7.2: Optimal contracts under complete and incomplete information scenarios.

The physical meaning of Theorem 1 is that the museum operator and the visitors provide rewards and contribution, respectively, in proportion to the visitors' types. The analysis presented in this section is graphically captured in Fig. 7.2.

7.5 Museum Congestion & Feedback Management Under Incomplete Information

In this section, we examine the general case, where the museum operator is unaware of the visitors' types $\tau_n, \forall n \in N$, and estimates them in a probabilistic manner.

7.5.1 Problem Formulation

Similarly as before, our goal is to determine the optimal contracts $\{r_n^*, \hat{x}_n^*\}, \forall n \in N$ between the museum operator and the visitors such that the museum operator optimizes its profit and benefits, while jointly satisfying the visitors' QoE prerequisites. To guarantee feasibility of the contracts, the following rational and necessary conditions should hold true: individual rationality (IR), incentive compatibility (IC), fairness, monotonicity, and rationality.

Definition 1. *(Individual Rationality (IR)) Each visitor, n, must experience a non-negative utility, i.e., $U_n(\hat{x}_n) = \tau_n e(r_n) - k\hat{x}_n \geq 0, \forall n \in N$, in the optimal contract, $\{r_n^*, \hat{x}_n^*\}$, in order to be incentivized to participate in the operation of the museum CPSS.*

Definition 2. *(Incentive Compatibility (IC)) An optimal personalized contract, $\{r_n^*, \hat{x}_n^*\}$, designed for a visitor of type τ_n should provide higher utility to the visitor compared to any other contract that is not aligned to the visitor's personal characteristics, i.e., $\tau_n e(r_n^*) - k\hat{x}_n^* \geq \tau_n e(r_{n'}) - k\hat{x}_{n'}, \forall n \neq n', n, n' \in N$.*

The physical meaning of the IR and IC conditions is that the contracts should be wisely designed to incentivize the visitors' participation in the museum's operation and should be personalized to each visitor's unique characteristics.

Theorem 2. *(Fairness) The optimal contract should be fair by assigning higher (or equal) reward to a visitor of higher (or equal) type, i.e.,* $r_n > r_{n'} \Leftrightarrow \tau_n > \tau_{n'} (r_n = r_{n'} \Leftrightarrow \tau_n = \tau_{n'})$.

Proof. Initially, we consider $\tau_n > \tau_{n'}, \forall n, n^{'} \in N$, with $n \neq n^{'}$. Based on the IC condition, we have:

$$\tau_n e(r_n) - k\hat{x}_n \geq \tau_n e(r_{n'}) - k\hat{x}_{n'} \tag{7.5}$$

$$\tau_n e(r_{n'}) - k\hat{x}_{n'} \geq \tau_{n'} e(r_n) - k\hat{x}_n \tag{7.6}$$

By adding the inequalities (7.5) and (7.6), we have:

$$(\tau_n - \tau_{n'})e(r_n) > (\tau_n - \tau_{n'})e(r_{n'}) \tag{7.7}$$

We know that $\tau_n > \tau_{n'}$, thus, we have $e(r_n) > e(r_{n'})$. Given that the evaluation function is strictly increasing with respect to r_n, we conclude that $r_n > r_{n'}$. We then consider $r_n > r_{n'}, \forall n, n^{'} \in N$, with $n \neq n^{'}$. Given the monotonicity of the evaluation function, we have $e(r_n) - e(r_{n'}) > 0$. Thus, based on Eq. 7.7, we have $(\tau_n - \tau_{n'})(e(r_n) - e(r_{n'})) > 0$. Therefore, we conclude that $\tau_n > \tau_{n'}$. □

The physical meaning of the fairness condition is that a visitor of higher type, who has the potential to contribute more in the museum's operation, should receive greater reward.

Theorem 3. *(Monotonicity) A visitor of higher type, i.e.,* $\tau_1 < \cdots < \tau_n < \cdots < \tau_{|N|}$ *receives greater reward, i.e.,* $r_1 < \cdots < r_n < \cdots < r_{|N|}$*, thus, they are expected to provide greater contribution, i.e.,* $\hat{x}_1 < \cdots < \hat{x}_n < \cdots < \hat{x}_{|N|}$.

Proof. Without loss of generality and for convenience in the notation, we consider that the visitors' types are sorted as $\tau_1 < \cdots < \tau_n < \cdots < \tau_{|N|}$. Then, based on Theorem 2, we can also show that $r_1 < \cdots < r_n < \cdots < r_{|N|}$. Thus, given that $r_n(\hat{x}_n) = \tau_n \hat{x}_n$, we can easily conclude that $\hat{x}_1 < \cdots < \hat{x}_n < \cdots < \hat{x}_{|N|}$. □

The physical meaning of Theorem 3 is that a visitor with higher type is capable of providing greater contribution and thus receives a greater reward from the museum operator.

Theorem 4. *(Rationality) Visitors of higher types, i.e.,* $\tau_1 < \cdots < \tau_n < \cdots < \tau_{|N|}$*, eventually achieve higher utilities, i.e.,* $U_1 < \cdots < U_n < \cdots < U_{|N|}$*.*

Proof. Based on the IC condition for two indicative visitors, $n, n^{'} \in N, n \neq n^{'}$, with $\tau_n > \tau_{n'}$, we have $\tau_n e(r_n) - k\hat{x}_n \geq \tau_n e(r_{n'}) - k\hat{x}_{n'}$. Because $\tau_n > \tau_{n'}$, we have $\tau_n e(r_n) - k\hat{x}_n \geq \tau_{n'} e(r_{n'}) - k\hat{x}_{n'}$. Generalizing the latter outcome for the visitors with types $\tau_1 < \cdots < \tau_n < \cdots < \tau_{|N|}$, we conclude that $U_1 < \cdots < U_n < \cdots < U_{|N|}$. □

The physical meaning of Theorem 4 is that a visitor of higher type, who provides greater contribution to the museum operation (Theorem 3), will receive greater reward (Theorem 2) and will thus eventually enjoy greater utility.

It should be highlighted that the above five conditions, as presented in Definitions 1, 2, and Theorems 2-4, are necessary but not sufficient in order to conclude to the optimal contract under the incomplete information scenario. Consequentially, we formulate the optimization problem to capture the interactions among the museum operator and the visitors as a maximization problem of the museum operator's utility (Eq. 7.8a) under the IR (Eq. 7.8b), IC (Eq. 7.8c), fairness, monotonicity, and rationality constraints, which can be jointly expressed in Eq. 7.8d, considering that $\tau_1 < \cdots < \tau_n < \cdots < \tau_{|N|}$. Therefore, the optimization problem of determining the optimal contracts under incomplete information of the visitors' types can be expressed as follows:

$$\max_{\{r_n,\hat{x}_n\}_{\forall n\in N}} U_M(\hat{\mathbf{x}}) = \sum_{n=1}^{|N|}[P_n(\hat{x}_n - cr_n)] \tag{7.8a}$$

$$\text{s.t.} \quad \tau_n e(r_n) - k\hat{x}_n \geq 0, \forall n \in N \quad \text{(IR)} \tag{7.8b}$$

$$\tau_n e(r_n) - k\hat{x}_n \geq \tau_n e(r_{n'}) - k\hat{x}_{n'}, \forall n \neq n', \in N \quad \text{(IC)} \tag{7.8c}$$

$$0 \leq r_1 < r_2 < \cdots < r_n < \cdots < r_{|N|} \tag{7.8d}$$

The optimization problem (7.8a)-(7.8d) is non-convex.

7.5.2 Problem Solution

To solve the optimization problem (7.8a)-(7.8d) and determine the optimal contracts, we reduce its constraints, as shown in the following analysis. Initially, we focus on the IR constraint (Eq. 7.8b). Based on Theorem 4 and the IC condition, we have that $\tau_n e(r_n) - k\hat{x}_n \geq \tau_n e(r_{n-1}) - k\hat{x}_{n-1} \geq \cdots \geq \tau_n e(r_1) - k\hat{x}_1$. We observe that if the IR condition holds true for the visitor of the lowest type, i.e., $\tau_n e(r_1) - k\hat{x}_1 \geq 0$, then it will hold true for any visitor of higher type. Furthermore, given that the museum operator will provide the just-sufficient reward to incentivize the visitors, we conclude that the constraint (7.8b) can be replaced by $\tau_n e(r_1) - k\hat{x}_1 = 0$.

We then analyze the IC constraint in Eq. 7.8c using the following terminology about the IC constraints of visitors with different types: (i) Downward IC (DIC) constraints for $n, n', n' \in \{1, \ldots, n-1\}$; (ii) Local DIC (LDIC) constraint for $n, n-1 \in N$; (iii) Upward IC (UIC) constraint for $n, n', n' \in \{n+1, \ldots, |N|\}$, and (iv) Local UIC (LUIC) constraint for $n, n+1 \in N$.

Theorem 5. *All of the UIC constraints can be equivalently captured by the LDIC constraint.*

Proof. We consider the IC conditions of three visitors, $n-1, n, n+1, \forall n \in N$, as follows:

$$\tau_{n-1}e(r_{n-1}) - k\hat{x}_{n-1} \geq \tau_{n-1}e(r_n) - k\hat{x}_n \tag{7.9}$$

$$\tau_n e(r_n) - k\hat{x}_n \geq \tau_n e(r_{n+1}) - k\hat{x}_{n+1} \tag{7.10}$$

Based on Theorem 2, Eq. 7.10 can be analyzed:

$$\begin{aligned} &k(\hat{x}_{n+1} - \hat{x}_n) \geq \tau_n[e(r_{n+1}) - e(r_n)] \xRightarrow{\tau_n > \tau_{n-1}} \\ &\qquad k(\hat{x}_{n+1} - \hat{x}_n) \geq \tau_{n-1}[e(r_{n+1}) - e(r_n)] \end{aligned} \tag{7.11}$$

Combining Eq. 7.9 and Eq. 7.11, we have $\tau_{n-1}e(r_{n-1}) - k\hat{x}_{n-1} \geq \tau_{n-1}e(r_n) - k\hat{x}_n \geq \tau_{n-1}e(r_{n+1}) - k\hat{x}_{n+1}$. By applying the latter outcome for all of the UIC constraints, we have $\tau_{n-1}e(r_{n-1}) - k\hat{x}_{n-1} \geq \tau_{n-1}e(r_n) - k\hat{x}_n \geq \tau_{n-1}e(r_{n+1}) - k\hat{x}_{n+1} \geq \cdots \geq \tau_{n-1}e(r_{|N|}) - k\hat{x}_{|N|}$. Thus, we conclude that all of the UIC constraints can be equivalently captured by the LDIC constraint, as expressed in Eq. 7.9. □

Theorem 6. *All of the DIC constraints can be equivalently captured by the LDIC constraint.*

Proof. We consider the IC conditions of three visitors, $n-1, n, n+1, \forall n \in N$, as follows:

$$\tau_{n+1}e(r_{n+1}) - k\hat{x}_{n+1} \geq \tau_{n+1}e(r_n) - k\hat{x}_n \tag{7.12}$$

$$\tau_n e(r_n) - k\hat{x}_n \geq \tau_n e(r_{n-1}) - k\hat{x}_{n-1} \tag{7.13}$$

Given that $\tau_n > \tau_{n-1}$, we have $r_n > r_{n-1} \overset{e\nearrow}{\Longleftrightarrow} e(r_n) - e(r_{n-1}) > 0$. Thus, we have $\tau_{n+1} > \tau_n \Leftrightarrow \tau_{n+1}[e(r_n) - e(r_{n-1})] > \tau_n[e(r_n) - e(r_{n-1})] \geq k(\hat{x}_n - \hat{x}_{n-1})$, where the last step holds true based on Eq. 7.13. We can then apply the latter outcome recursively for all of the LDIC constraints, as in $\tau_{n+1}e(r_{n+1}) - k\hat{x}_{n+1} \geq \tau_{n+1}e(r_n) - k\hat{x}_n \geq \tau_{n+1}e(r_{n-1}) - k\hat{x}_{n-1} \geq \cdots \geq \tau_{n+1}e(r_1) - k\hat{x}_1$. Therefore, we conclude that all of the DIC constraints can equivalently be captured by the LDIC constraint, as expressed in Eq. 7.13. □

Combining the outcomes of Theorem 5 and 6, we conclude that all of the IC constraints can be reduced to the LDIC constraint, i.e., $\tau_n e(r_n) - k\hat{x}_n \geq \tau_n e(r_{n-1}) - k\hat{x}_{n-1}$. By considering the IR and IC constraints' reductions, the optimization problem (7.8a)-(7.8d) can be written as follows:

$$\max_{\{r_n, \hat{x}_n\}_{\forall n \in N}} U_M(\hat{\mathbf{x}}) = \sum_{n=1}^{|N|} [P_n(\hat{x}_n - cr_n)] \tag{7.14a}$$

$$\text{s.t.} \quad \tau_n e(r_1) - k\hat{x}_1 = 0 \quad \text{(IR)} \tag{7.14b}$$

$$\tau_n e(r_n) - k\hat{x}_n \geq \tau_n e(r_{n-1}) - k\hat{x}_{n-1} \quad \text{(LDIC)} \tag{7.14c}$$

$$0 \leq r_1 < r_2 < \cdots < r_n < \cdots < r_{|N|} \tag{7.14d}$$

The optimization problem (7.14a)-(7.14d) is convex and can be solved by standard convex optimization tools. The outcome is the optimal contracts $\{r_n^*, \hat{x}_n^*\}, \forall n \in N$. Thus, the museum operator determines the optimal allocated rewards $r_n^*, \forall n \in N$ to the visitors, while the visitors decide their optimal contribution $\hat{x}_n^*, \forall n \in N$, consisting of the ratio of the number of provided evaluations about the exhibits and their time invested in the museum touring. The analysis presented in this section is graphically captured in Fig. 7.2.

7.6 Evaluation & Results

In this section, a detailed set of numerical results are presented in order evaluate the performance of the proposed framework and reveal its operation benefits. In particular, the pure operation characteristics and performance of the proposed framework is presented in Section 7.6.1 considering both the complete and incomplete information scenarios, while the impact of the visitors personal cost to provide their contribution on the overall museum CPSS operation is discussed in Section 7.6.2. The benefits of addressing the visitors in a personalized manner, while considering their unique characteristics, are shown in Section 7.6.3. Finally, a thorough comparative evaluation is demonstrated in Section 7.6.4, which considers different alternative scenarios of visitors' invested contribution and their impact on both their satisfaction and the overall social welfare of the CPSS.

For the evaluation purposes, we have simulated a large size museum, considering $E_n \in [0, 380, 000]$, $t_n \in [1, 360]$ min, and t_n^{Max} following a random and uniform distribution in $[60, 360] min$. Also, we assume that p_n, P_n, and k_n are uniformly distributed in $(0, 1]$. Finally, we consider $|N| = 100$ visitors, and $c = 2, k = 5$, unless otherwise explicitly stated.

7.6.1 Pure Operation Performance

In this section, we examine the pure operation performance of the proposed museum and visitors interaction and feedback orchestration mechanism, with consideration of both the benchmarking scenario which features complete information, and the realistic scenario which features incomplete information regarding the visitors' types.

Fig. 7.3a shows the visitors' types, τ_n, as a function of their index (ID), where the visitors have been sorted with respect to an increasing value of their type. Similarly,

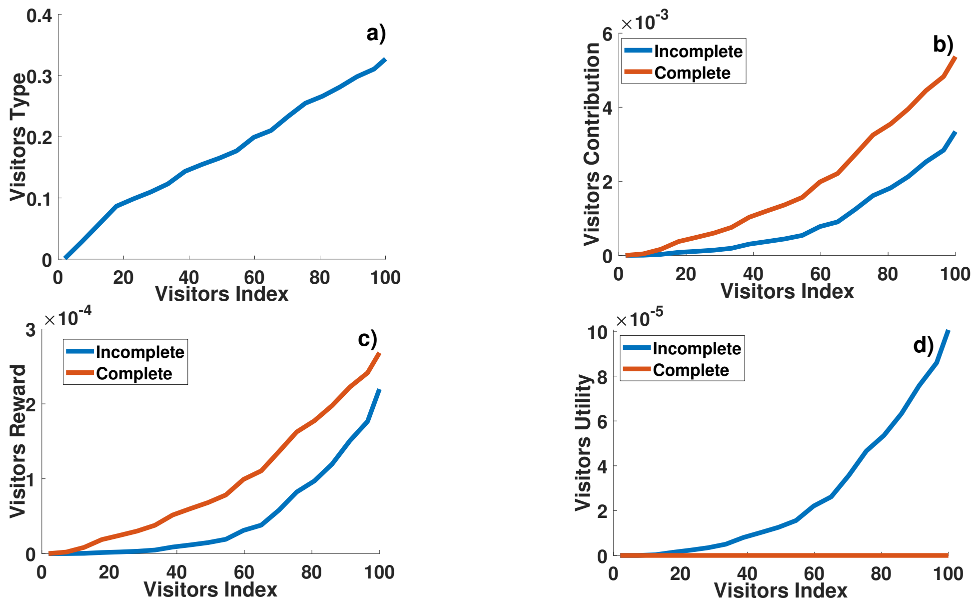

Figure 7.3: Visitors' (a) Type, (b) Contribution, (c) Reward, and (d) Utility as a function of their ID under the complete and incomplete information scenarios.

Fig. 7.3b-7.3d illustrate the visitors' normalized contribution $\hat{x}_n$, their received reward r_n, and their corresponding achieved utility $U_n(\hat{x}_n)$, respectively, as a function of their ID under the complete and incomplete information scenarios. The results reveal that the visitors of greater type provide greater contribution (Fig. 7.3b) to the operation of the museum and thereby receive a greater reward (Fig. 7.3c), following the fairness (Theorem 2) and monotonicity (Theorem 3) properties. Accordingly, the visitors of higher type eventually enjoy a greater utility (Fig. 7.3d) based on the rationality property (Theorem 4). Moreover, focusing on the comparison of the complete and incomplete information scenarios, we observe that the museum operator can fully exploit the visitors' capabilities in collecting their contributions. This is derived from the fact that the visitors provide a greater contribution under the complete information scenario (Fig. 7.3b), and are thus rewarded more than in the incomplete information scenario (Fig. 7.3c), while achieving their minimum-acceptable utility (i.e., zero) which is just sufficient enough to incentivize their participation in the overall process (Fig. 7.3d). The latter outcome is important to the museum op-

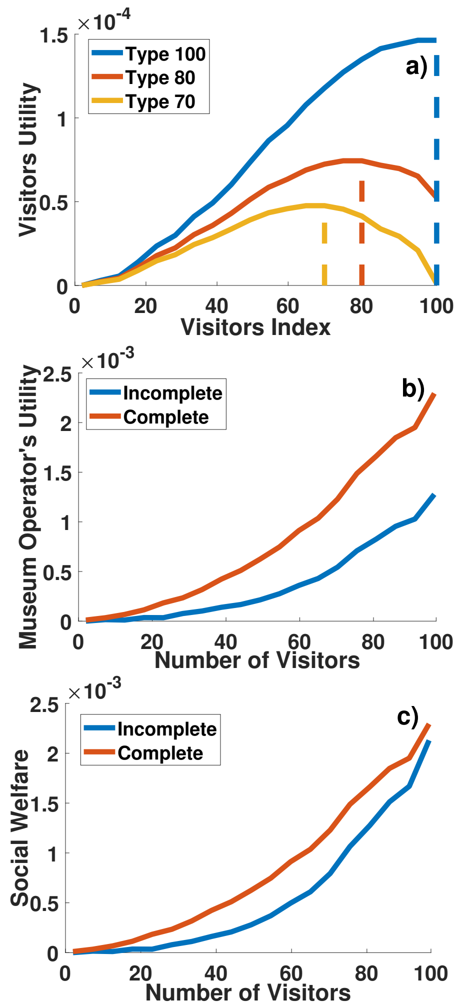

Figure 7.4: (a) Visitors' incentive compatibility, (b) Museum operator's cumulative utility, and (c) Social welfare.

erator, who strives to wisely invest their resources such that valuable information is collected from the visitors, while the latter ones spend wisely their time in the museum touring.

Fig. 7.4a depicts the achieved utility for three indicative visitors' types, i.e., types $70, 80$, and 100, considering the optimal contracts that were determined under the incomplete information scenario for each corresponding visitor's index, i.e., $\{r_n^*, \hat{x}_n^*\}, \forall n \in N$. The results confirm that a visitor will achieve the highest relative utility when they receive a contract that has been designed with their own unique type and characteristics in consideration. For example, the visitor of type 80 will receive the maximum utility for the optimal personalized contract $\{r_{80}^*, \hat{x}_{80}^*\}$, while lower utility is achieved for any other optimal contract $\{r_n^*, \hat{x}_n^*\}, n \neq 80, \forall n \in N$ that is not aligned to its personal type τ_{80}. The following observation respects and confirms the incentive compatibility property (Definition 2) and shows the importance of treating the visitors in a personalized manner. Furthermore, Fig. 7.4a shows that the visitors of higher type will experience higher utility, when receiving their personalized contract. The latter observation derives from the rationality condition, as presented in Theorem 4. Moreover, Fig. 7.4b and Fig. 7.4c illustrate the museum operator's cumulative utility and the overall CPSS's social welfare as a function of the number of visitors. The results reveal that, as expected, better outcomes can be achieved under the complete information scenario. However, it should be noted that the performance achieved in the realistic scenario of incomplete information is approximately 35% less than in the ideal case of complete information, thereby showing that the proposed framework is robust and efficient under the uncertainties introduced by realistic conditions.

7.6.2 Impact of Visitor's Personal Cost

In this section, we examine the impact of the visitor's personal cost $k, k \in \mathbb{R}^+$ on the interaction among the visitors and the museum operator. The visitor's personal cost may stem from various factors, such as the visitor's smartphone's battery drainage during the exhibit evaluations via the museum's mobile application, delay in touring in order to provide evaluations about the exhibits, and others. Three different comparative scenarios are considered capturing a (i) high ($k = 5$), (ii) medium ($k = 3$), and (iii) low ($k = 2$) personal cost.

Fig. 7.5a-7.5c show the visitor's utility as a function of the visitor's ID, the

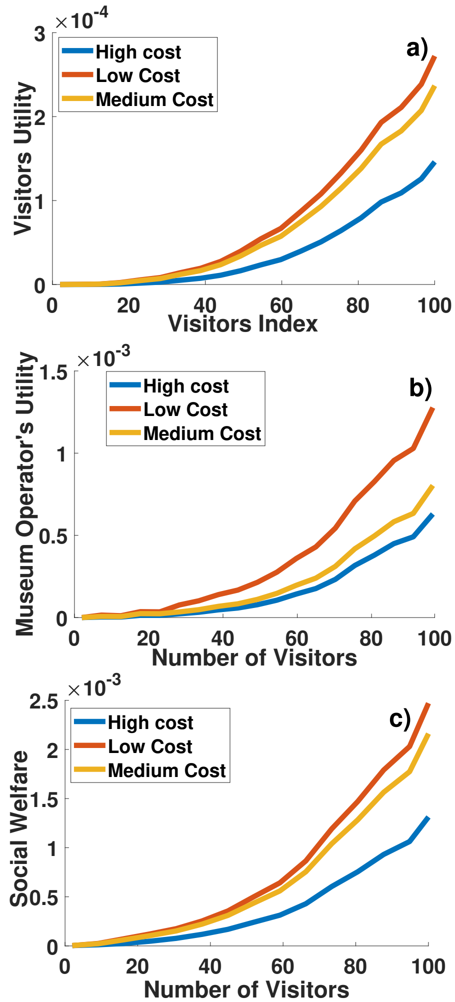

Figure 7.5: Impact of the visitor's personal cost on the system's operation.

museum operator's utility, and the overall social welfare as a function of the number of visitors, respectively. The results demonstrate that the higher the visitor's personal cost is, the less utility they enjoy (Fig. 7.5a). Moreover, given the hesitant behavior of the visitors to provide their contribution when they experience higher personal cost, the museum operator's utility (Fig. 7.5b) and the overall social welfare (Fig. 7.5c) achieve low levels as well.

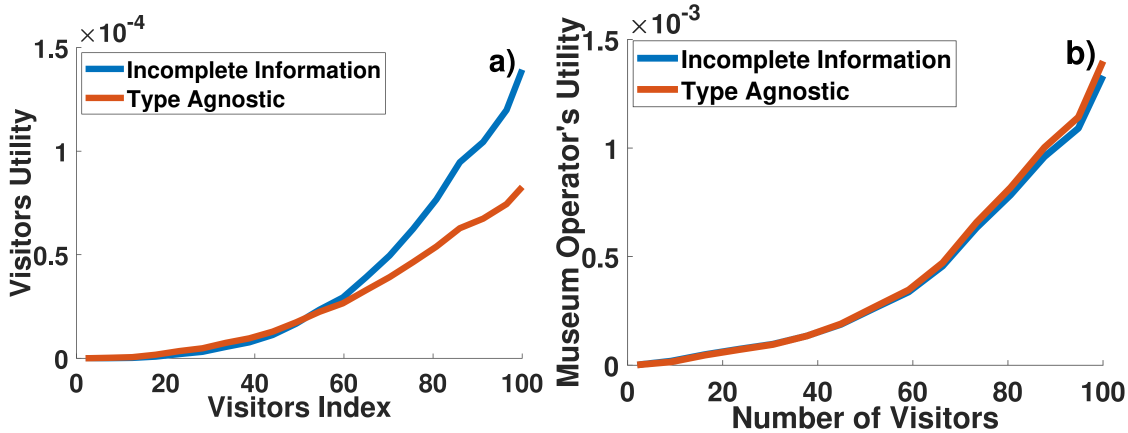

Figure 7.6: Labor economics-based vs type-agnostic interaction and feedback orchestration.

7.6.3 Visitors' Unique Types and Characteristics

In this section, we examine the impact of considering the visitors' unique types and characteristics, in the overall proposed framework of museum and visitors interaction and feedback orchestration. In particular, we consider the proposed labor economics-based approach under the realistic incomplete information scenario, and compare it against a type-agnostic scenario, where the rewards are allocated to the visitors in a homogeneous manner, i.e., $r_n(\hat{x}_n) = \frac{\sum_{n=1}^{|N|} \tau_n}{|N|} \hat{x}_n$.

Fig. 7.6a-7.6b illustrate the visitors' utility with respect to their ID, and the museum operator's cumulative utility as a function of the number of visitors, respectively, considering both the incomplete information and the type-agnostic scenarios. The results reveal that the visitors achieve, on average, 25% better utility (Fig. 7.6a) under the proposed labor economics-based approach when compared to the type-agnostic approach. This outcome stems from the personalized treatment of the visitors by the museum operator in our proposed framework, where the museum operator provides rewards well-aligned to the visitors' characteristics. Also, it is observable that the type-agnostic scenario over-rewards visitors of lower types and under-rewards visitors of higher types, thereby making the rewards allocation less fair. Moreover, based on the results presented in Fig. 7.6b, we conclude that the proposed labor economics-based scheme achieves to also outperform the type-agnostic one from the overall system perspective (i.e., museum operator's utility). Specifically a 4% improvement is observed for the case 100 visitors while, based on observed trends, it is evident that an even greater improvement upon type-agnosticism could be achieved for cases with larger populations of visitors.

7.6.4 Comparative Evaluation

In this section, a thorough comparative evaluation is provided considering various scenarios regarding the visitors' contribution $\hat{x}_n$. In addition to our proposed framework, four different alternatives are considered: (i) Minimum contribution, i.e., $\hat{x}_n = \min\{\hat{x}_n\}_{\forall n \in N}$, (ii) Maximum contribution, i.e., $\hat{x}_n = \max\{\hat{x}_n\}_{\forall n \in N}$, (iii) Random contribution, and (iv) Guided contribution, i.e., $\hat{x}_n = f(\hat{x}_n^*) = \frac{-log(\tau_n)}{100}$.

Fig. 7.7a shows the visitors' utility with respect to their ID, while Fig. 7.7b-7.7c present the museum operator's utility, and the CPSS's social welfare, respectively, as a function of the number of visitors. The results reveal that the visitors' maximum contribution scenario benefits the museum operator but causes visitor satisfaction to be low and, thus, yields low social welfare. The exact opposite holds true for the minimum contribution scenario, wherein the visitors are not engaged with the museum's operation. The random and guided contribution scenarios present an intermediate performance for both the visitors' and the museum operator's benefits as compared to the aforementioned extreme scenarios of maximum and minimum contribution. Finally, it is observed that our proposed framework benefits the visitors in terms of improving their perceived satisfaction, while also supporting the museum operator's needs to collect feedback from the visitors and incentivize them to wisely spend their

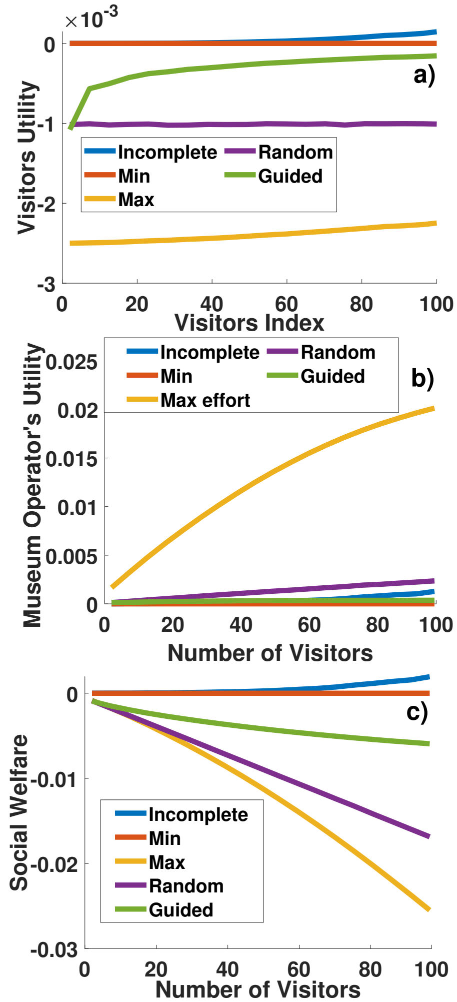

Figure 7.7: Comparative evaluation.

time during their museum touring. The combined benefit of our proposed framework is depicted in the superior social welfare achieved, which shows a five-fold improvement compared to the maximum contribution scenario, where the visitors provide the maximum possible number of evaluations and invest the minimum acceptable time for their tour.

7.7 Conclusion and Future Work

In this chapter, the problem of jointly orchestrating the museum operator's and the visitors' interaction, as well as the feedback provided by the visitors while accounting for their behavioral characteristics, is considered. The problem is treated and solved under the prism and reasoning of a labor-economics-based approach. In particular, the museum is treated as a cyber-physical-social system and the visitors' unique characteristics are derived to define their unique types. Following the principles of labor economics, an optimization problem is formulated and solved to jointly determine the visitors' optimal contributions in the museum's operation and the optimal rewards allocated by the museum operator to incentivize the visitors' engagement. The scenarios of both complete and incomplete information use cases regarding the visitors' characteristics are examined, for benchmarking and realistic implementations purposes, respectively. A set of detailed simulations considering a large size museum, e.g., Louvre Museum, is provided to demonstrate the performance and benefits of the proposed framework.

Part of our current and future work refers to the actual implementation of the proposed framework in a mobile Android and iOS application, and its pilot testing at the Acropolis Museum in Athens, Greece.

Chapter 8

Prosumer-centric Self-sustained Smart Grid Systems

8.1 Introduction

Smart Grid (SG) systems have been introduced as an alternative solution to traditional power systems which operate in a centralized manner, generating power in large power stations via the exploitation of fossil fuel resources, and distributing the generated power to consumers [153]. SGs consist of multiple microgrids, which are small-scale power supply networks, accommodating conventional energy units, renewable energy sources, and energy storage systems [154]. One key enabler of SGs is the new type of users, named *prosumers*, who are able to generate, store, sell, and buy energy by mainly exploiting solar photovoltaic panels and storage devices [155], among others. The prosumers are equipped with smart meters to exchange (sell/buy) power with the Microgrid Operator (MGO), thus creating a local energy trading system [156]. In this chapter, we capture the prosumers' interactions with the MGO in terms of selling and buying power based on a labor economics framework, while guaranteeing the joint optimization of their profits and enabling the overall microgrid system to converge to a stable point of operation.

8.1.1 Related Work

Several recent research works focus on the operation of microgrid systems aiming to satisfy the consumers' or prosumers' power demand via dealing with the Demand Response Management (DRM) problem [157]. In [158], the interactions among multiple microgrid systems are studied based on the Nash bargaining theory in order to incentivize each microgrid to participate in the proactive energy trading and fair benefit sharing. The authors formulate the corresponding joint optimization problem and solve it by decomposing the problem into two sequential problems, where the first minimizes the social cost and the second one optimizes the trading benefit sharing. The problem of high-levels stochasticity in the energy production of the renewable energy sources is studied in [159]. The authors provide a systematic approach to deal with this problem and provide the enhanced flexibility to the system to satisfy the consumers' power demand via exploiting the fast-ramping units, the energy storage, and the hourly demand response. In [160], the authors introduce a novel transactive energy control mechanism and a pricing rule to capture the interactions among multiple microgrids, aiming at jointly minimizing their operating cost and optimizing the utilization of the renewable energy sources. The authors have provided a detailed comparative evaluation to other centralized and decentralized transactive energy control mechanisms to show the benefits of the proposed approach in terms of the microgrids' effective operation and computational efficiency in microgrids' coordination.

The problem of reducing prosumers' electricity bills, while guaranteeing their minimum power demand constraints is studied in [161] via the introduction of an intelligent residential energy management system. A predictive mechanism of the power demand and supply in a microgrid is introduced in [162] by designing a smart load estimator based on a neural networks' approach. The designed mechanism considers the ambient temperature, the time of day, the hourly price, and the peak demand. It should be noted here that the above research works follow a system-based approach emphasizing on the operation of the microgrid, without however accounting for the unique and personal characteristics of the prosumers.

Focusing on the prosumer-centric microgrid systems, a prospect-theoretic energy trading approach is introduced in [163], in which the prosumers' risk-aware characteristics are considered based on the uncertainty that the selling/buying energy price introduces. The authors formulate a single-leader multiple-follower Stackelberg game, where the microgrid operator (leader) announces the optimal price and the prosumers (followers) determine the amount of energy that they sell or buy, with all the involved entities aiming to optimize their profit. A similar Stackelberg-based approach is followed in [18] that also introduces a reinforcement learning mechanism to enable the consumers to select the utility company that they will purchase energy from, in an autonomous manner. In [164], the prosumers consider the energy as a heterogeneous product depending on the generation technology, its location in the SG, and its owner's reputation. Accordingly, an optimization problem is formulated to minimize the costs of energy losses and battery depreciation, while accounting for the prosumers' preferences regarding the energy.

Placing further emphasis on exploiting the prosumers' unique power generation and demand characteristics, a pricing-based DRM problem is introduced in [165], which jointly considers the prosumers' behavioral characteristics in terms of consuming electricity and the electricity demand of their household devices, which can be of various types. In [166], the authors introduce a distributed system-wide framework aimed at minimizing the prosumers' payments, while guaranteeing their privacy and comfort constraints, via dynamically adapting the system load profile. In [167, 168], the authors study the impact of the communication unreliability among the MGO and the prosumers on the DRM performance and the electricity price by formulating a joint maximization problem of the DRM performance with respect to the electricity consumption and price, and solve it by leveraging the dual decomposition method. Labor economics and Contract Theory have been also introduced in the literature in order to incentivize the prosumers to follow a desired behavior within a microgrid [39]. In particular, in [20], the authors introduce a labor economics framework to capture the interactions of the prosumers and the MGO. A contract-theoretic optimization problem is formulated and solved to determine the optimal amount of purchased electricity and the optimal rewards provided by the MGO to the prosumers for the sake of both parties having optimized profit. It is highlighted that this research work considers the prosumers only as buyers and *not* sellers.

8.1.2 Contributions and Outline

Despite the efforts made in the previous research works, in regards to system-centric or prosumer-centric operation of microgrids, how to incorporate the dual role of the prosumer, i.e., seller and buyer, within the operation of the microgrid system still remains an open issue. Furthermore, within such a setting, facilitating the smooth and seamless operation of the microgrid system, while incentivizing the prosumers to act in a desirable manner and simultaneously considering their unique personal energy generation and demand characteristics is even more challenging.

In this research work, we strive exactly to tackle these issues by introducing a contract-theoretic framework to capture the interactions of the prosumers, acting either as sellers or buyers, with the microgrid [39]. In particular, a labor economics-based approach is designed and evaluated, exploiting the principles of Contract Theory to jointly achieve the satisfaction of the various system entities, that is the MGO and the prosumers, which often present competing interests. Accordingly, the relations between the MGO and the prosumers (sellers or byers) are captured following the model of employer-employee relationship, while aiming to jointly satisfy the profit and requirements of the involved competing parties. Specifically, the main contributions of this research work that differentiate it from the rest of the existing literature are summarized below.

1. A microgrid system consisting of the microgrid operator (MGO) and the prosumers, who generate energy based on renewable energy sources (e.g., solar photovoltaic panels) and are equipped with energy storage (e.g., Lithium-ion batteries) is introduced. Within the considered microgrid, the prosumers can dynamically act as sellers or buyers based on their energy generation, demand, and storage characteristics over time.

2. The interactions among the sellers and the MGO are captured via a contract-theoretic optimization problem which determines the optimal amount of energy that the sellers sell to the MGO at a specific announced price, and the optimal rewards (e.g., price discount) offered by the MGO. The goal of the formulated problem is to maximize the MGO's profit, while jointly optimizing the profit of the sellers via considering their unique personal energy generation, demand, and storage characteristics.

3. Focusing on the buyers' side, a different contract-theoretic optimization problem is formulated to study the interactions among the buyers and the MGO. The MGO provides personalized rewards to the buyers, e.g., fixed price, considering their energy demand, while the buyers invest their "effort", i.e., money, to purchase the amount of energy that covers their demand. The optimal personalized contracts, i.e., optimal personalized reward and purchased energy per

buyer, are determined to bring the dynamic interaction of the MGO and the prosumers into a stable mode of operation.

4. A detailed series of experiments are performed to show the drawbacks and benefits of the proposed prosumer-centric self-sustained smart grid system's operation approach. This is realized under both a benchmarking scenario of complete information and a realistic scenario of incomplete information regarding the prosumers' energy generation, demand, and storage characteristics. A scalability analysis is performed to show the efficiency and robustness of the proposed framework. Also, a detailed study is performed regarding the impact of the prosumers' energy generation and demand characteristics, as well as the MGO's pricing policies, on the interactions of the sellers and buyers with the MGO.

The rest of the chapter is organized as follows. Section 8.2 introduces the human-centric smart grid system model, while Sections 8.3 and 8.4 introduce and solve the contract-theoretic optimization problems for the sellers and buyers, respectively. Simulation and comparative results are presented in Section 8.5. Finally, Section 8.6 concludes the chapter.

8.2 Human-centric Smart Grid System Model

A microgrid system is considered, consisting of the microgrid operator (MGO) and the prosumers. The prosumers can generate energy via various alternative options, such as solar photovoltaic panels and small wind turbine power generation systems, and can also store the energy in storage systems, such as Lithium-ion batteries [169]. Each prosumer's residential infrastructure is equipped with a smart meter to dynamically measure the energy generation, demand, storage, and exchange (selling or buying) power with the MGO. We examine the interactions of the MGO and the prosumers at each time slot t, with the set of times slots denoted as $T = \{1, \dots, t, \dots, |T|\}$. The sets of prosumers, sellers, and buyers are denoted as $N = \{1, \dots, n, \dots, |N|\}$, $S = \{1, \dots, s, \dots, |S|\}$, and $B = \{1, \dots, \beta, \dots, |B|\}$, respectively, with $S \subseteq N$, $B \subseteq N$, and $|S| + |B| = |N|$.

Each prosumer has a set of home appliances, $A_n = \{1, \dots, a_n, \dots, |A_n|\}$, which at the duration of one time slot t (e.g., one hour) can be either on, i.e., $\delta^t_{a_n} = 1$, or off, i.e., $\delta^t_{a_n} = 0$. Thus, their total energy demand is $d^t_n = \sum_{\forall a_n \in A_n} \delta^t_{a_n} E_{a_n}$ [kWh] in the duration of a time slot t, where E_{a_n} [KWh] is the energy consumption of the appliance a_n when it is operating during time slot t [170]. It should be noted that a prosumer can shift the operation of some appliances over time, thus, $d^{Min}_n \leq d^t_n \leq d^{Max}_n$. Specifically, d^{Min}_n [kWh] denotes the total energy demand of the appliances of the prosumer n that are non-shiftable over time, e.g., refrigerator or alarm system, while d^{Max}_n [kWh] captures the maximum possible energy demand if all of the prosumer's appliances are active. Thus, the prosumers' energy demand vector is defined as $\mathbf{D} = [d^t_1, \dots, d^t_n, \dots, d^t_{|N|}]$ per time slot t. Also, the prosumers can generate energy by exploiting their own renewable energy sources. Thus, the prosumers' renewable energy generation vector is defined accordingly as $\mathbf{G} = [g^t_1, \dots, g^t_n, \dots, g^t_{|N|}]$ [kWh] per time slot t. The prosumers can act either as sellers or buyers per time slot t based on their personal energy generation and demand characteristics. Thus, in the following analysis, we examine the sellers and buyers cases.

Sellers Case: If $g^t_n + b^{t-1}_n \geq d^t_n$, the prosumer can cover their energy demand without purchasing energy from the MGO, while also dynamically deciding to sell the energy generation surplus to the MGO. The energy generation surplus is calculated as $b^{t+1}_n = b^t_n + (g^t_n - d^t_n)$ [kWh], where a percentage $e^t_s \in [0, 1]$ of it can be sold to the MGO. The energy generation surplus is assumed to be stored in the prosumer's energy storage system, e.g., Lithium-ion batteries. In this case, the prosumer n acts as a seller. The sellers can be incentivized by the MGO to sell their energy surplus into the energy market instead of storing it locally for future use, if appropriately designed personlized rewards, such as fixed energy price, are provided by the MGO.

Buyers Case: If $g_n^t + b_n^{t-1} < d_n^t$, the prosumer's total generated and stored energy is not sufficient to cover their energy demand d_n^t. Thus, the prosumer n acts as a buyer β, and aims to purchase $d_n^t - g_n^t - b_n^{t-1}$ [kWh] amount of energy from the MGO, in consideration of their personal energy needs which are shaped by their respective shiftable and non-shiftable demands.

In the subsequent two sections, we study the overall interactions of the sellers and the buyers with the MGO, in terms of selling a purchasing energy, with consideration of their unique personal energy generation and demand characteristics. The architecture of the overall prosumer-centric self-sustained smart grid system is presented in Fig. 8.1.

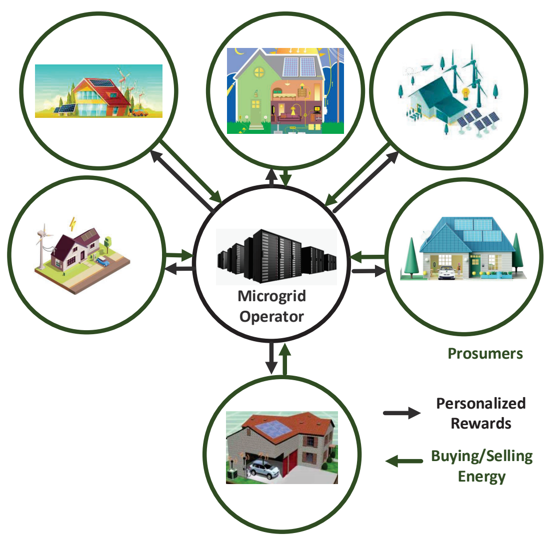

Figure 8.1: Prosumer-centric self-sustained smart grid system.

8.3 Sellers' & Electricity Market's Interactions

In this section, we capture the interactions of the sellers with the MGO in terms of determining the optimal amount of energy that they sell based on the appropriate incentives, i.e., rewards, provided by the MGO. Each seller $s, \forall s \in S \subseteq N$ is characterized by their type $\tau_s^t = \frac{b_s^{t+1}}{\sum_{\forall s \epsilon S} b_s^{t+1}} \in [0,1]$, which represents their normalized energy surplus among the sellers, thus, showing their potential to sell energy to the MGO. For notation convenience in the presentation, we consider $\tau_1^t < \cdots < \tau_s^t < \cdots < \tau_{|S|}^t$.

Based on the principles of Contract Theory [39], each seller acts as an "employee" investing their personal effort to the "employer", i.e., MGO, while the MGO incentivizes the sellers by providing personalized rewards, e.g., fixed energy price, in order for the MGO and the sellers to jointly optimize their achieved utility. The sellers' and the MGO's utility functions, as defined below in Eq. 8.1 and Eq. 8.2, respectively, represent their actual profit (i.e., satisfaction). The seller's effort is defined as $e_s^t = \frac{b_s^{t+1}}{\max_{\forall s \epsilon S}\{b_s^{t+1}\}} \in [0,1]$, showing the relative capability of each seller to sell energy to the MGO. Given that $\tau_1^t < \cdots < \tau_s^t < \ldots \tau_{|S|}^t$, we have $e_1^t < \cdots < e_s^t < \ldots e_{|S|}^t$. The MGO provides personalized rewards $r_s^t = \tau_s^t e_s^t$ to the sellers to incentivize them to sell their available energy surplus. Therefore, a seller with a higher potential to sell energy, who indeed sells a large amount of energy, will receive a high reward. The interaction among the MGO and the sellers, aiming at the joint optimization of their profit by participating in the energy market, concludes to an optimal contract (e_s^{t*}, r_s^{t*}) consisting of the optimal seller's effort e_s^{t*} and the MGO's optimal provided personalized reward r_s^{t*}.

Based on the above discussion, the utility functions of the sellers and the MGO are designed as the actual profit of the participants, while interacting among each other in the microgrid. The seller's utility is captured by the received revenue from selling energy to the MGO (first term of Eq. 8.1) while also considering their personal cost to locally produce the energy via the exploitation of their personal renewable energy source infrastructure (second term of Eq. 8.1).

$$U_s^t(e_s^t) = \tau_s^t \epsilon(e_s^t) - p_{\mathcal{S}} e_s^t \tag{8.1}$$

The seller's personal cost to produce their energy locally is denoted as $p_{\mathcal{S}} \in \mathbb{R}^+$ and, in the current analysis, is assumed to be a unitless number. This parameter can be mapped to monetary units, i.e., [$/kWh], when transferring this model in a real-life implementation and business case. Also, the function $\epsilon(e_s^t)$ represents the evaluation function, i.e., the way a seller interprets the received reward as personal satisfaction based on the enjoyed revenue. The evaluation function is a strictly increasing, concave, and continuous function with respect to the received reward, as a seller satisfaction increases monotonically with respect to the received reward, while at some point, the seller's satisfaction becomes saturated. For demonstration purposes, and without loss of generality, we consider $\epsilon(e_s^t) = \sqrt{r_s^t(e_s^t)}$.

The MGO's utility from interacting with the sellers is defined as follows:

$$U_{MGO,buy}^t(\mathbf{e}) = \sum_{s=1}^{|S|} Pr_s^t[e_s^t - r_s^t(e_s^t)] \tag{8.2}$$

where $\mathbf{e} = [e_1^t, \ldots, e_s^t, \ldots, e_{|S|}^t]$ is the sellers' effort vector. In general, the MGO is unaware of the sellers' energy generation demand and storage characteristics, which define the sellers' types. Thus, the MGO estimates each seller's type τ_s^t with probabilility Pr_s^t, where $\sum_{s=1}^{|S|} Pr_s^t = 1$. Several types of probability distributions, such as Gaussian, Poisson, and others, can be adopted based on the nature of the examined energy market. The specific distributions of the seller's types, and the corresponding probabilities Pr_s^t, can be determined in a real-life scenario according to the prosumers' energy characteristics that can be collected from their monthly electricity bills, based on a statistical or machine learning analysis. Note that Eq. 8.2 represents the MGO's profit from buying energy from the sellers via their invested effort e_s^t (first term of Eq. 8.2), while considering the MGO's cost to provide rewards r_s^t to the sellers (second term of Eq. 8.2).

8.3.1 Complete Information Scenario

Initially, we consider the benchmarking scenario, where the MGO has *complete information* about the sellers' types. The MGO aims to maximize its own profit from each seller of known type τ_s^t (Eq. 8.3a), while providing sufficient rewards to the sellers to maintain their business interactions and energy surplus sales (Eq. 8.3b). Thus, under the complete information scenario regarding the sellers' types, the interactions between the MGO and the sellers can be captured by the following contract-theoretic optimization problem:

$$\max_{\{e_s^t\}_{\forall s \in S}} [e_s^t - r_s^t(e_s^t)] \tag{8.3a}$$

$$\text{s.t.} \quad \tau_s^t \epsilon(e_s^t) - p_{\mathcal{S}} e_s^t \geq 0, \forall s \in S \tag{8.3b}$$

In this case, given that the MGO will provide just the sufficient rewards to incentivize the sellers to sell their energy surplus, Eq. 8.3b can be considered as an equality.

Theorem 1. *The optimal personalized contract between the MGO and each seller under the complete information scenario is* $(e_s^{t*}, r_s^{t*}) = (\frac{{\tau_s^t}^2}{2p_{\mathcal{S}}^2}, \frac{{\tau_s^t}^2}{4p_{\mathcal{S}}^2})$.

Proof. By solving Eq. 8.3b as an equality with respect to the reward, we have $r_s^t = (\frac{p_{\mathcal{S}} e_s^t}{\tau_s^t})^2$. By substituting the latter outcome in Eq. 8.3a, taking the first order derivative with respect to the effort e_s^t, and setting the outcome equal to zero, we conclude that $e_s^t = \frac{{\tau_s^t}^2}{2p_{\mathcal{S}}^2}$. Thus, the optimal contract is $(e_s^{t*}, r_s^{t*}) = (\frac{{\tau_s^t}^2}{2p_{\mathcal{S}}^2}, \frac{{\tau_s^t}^2}{4p_{\mathcal{S}}^2})$. □

The above outcome can be used mainly for benchmarking purposes, as sellers will not reveal their private information regarding their types, i.e., energy surplus, to the MGO, in a real-life scenario.

8.3.2 Incomplete Information Scenario

In the remaining analysis of this section, we examine the *incomplete information* scenario regarding the sellers' types. In pursuit of capturing the interactions between the sellers and the MGO, five fundamental conditions are examined: individual rationality (IR), incentive compatibility (IC), fairness, monotonicity, and rationality. Those conditions are necessary and sufficient in order to guarantee the feasibility and existence of an optimal contract among the MGO and the sellers. Each condition is analyzed and proved below, while its physical meaning is provided within the context of the MGO's and the sellers' interaction.

Definition 1. *(Individual Rationality (IR)) Each seller should receive a non-negative utility, i.e., $U_s^t(e_s^t) = \tau_s^t \epsilon(e_s^t) - p_{\mathcal{S}} e_s^t \geq 0, \forall s \in S$, from the optimal contract (e_s^{t*}, r_s^{t*}).*

Definition 2. *(Incentive Compatibility (IC)) Each seller achieves the maximum possible utility when they receive a contract aligned with their personal energy generation, demand, and storage characteristics, i.e., $\tau_s^t \epsilon(e_s^t) - p_{\mathcal{S}} e_s^t \geq \tau_s^t \epsilon(e_{s'}^t) - p_{\mathcal{S}} e_{s'}^t, \forall s, s' \in S$.*

The physical meaning of the IR and IC conditions is that each seller should be appropriately incentivized by the MGO by enjoying a positive profit aligned with their personal characteristics in order to sell their energy in the microgrid.

Proposition 1. *(Fairness) An optimal contract is fair, i.e., a seller of higher (or equal) type should enjoy a higher (or equal) reward: $r_s^t > r_{s'}^t \Leftrightarrow \tau_s^t > \tau_{s'}^t (r_s^t = r_{s'}^t \Leftrightarrow \tau_s^t = \tau_{s'}^t)$.*

Proof. We prove the sufficiency and necessity of the fairness condition. Assuming that $\tau_s^t > \tau_{s'}^t$, we can write the following IC constraints for the sellers s, s', $\forall s,\ s' \epsilon S$, $s \neq s'$.

$$\tau_s^t \epsilon(e_s^t) - p_{\mathcal{S}} e_s^t \geq \tau_s^t \epsilon(e_{s'}^t) - p_{\mathcal{S}} e_{s'}^t \tag{8.4}$$

$$\tau_{s'}^t \epsilon(e_{s'}^t) - p_{\mathcal{S}} e_{s'}^t \geq \tau_{s'}^t \epsilon(e_s^t) - p_{\mathcal{S}} e_s^t \tag{8.5}$$

By adding Eq. 8.4 and Eq. 8.5, we have:

$$(\tau_s^t - \tau_{s'}^t)\epsilon(e_s^t) \geq (\tau_s^t - \tau_{s'}^t)\epsilon(e_{s'}^t) \tag{8.6}$$

We know that $\tau_s^t > \tau_{s'}^t$, and $\epsilon(r_s^t(e_s^t))$ is a strictly increasing function with respect to r_s^t, thus we conclude that $r_s^t > r_{s'}^t$.

On the other hand, assuming that $r_s^t > r_{s'}^t$, we derive that $\epsilon(r_s^t(e_s^t)) > \epsilon(r_{s'}^t(e_{s'}^t))$, thus we rewrite Eq. 8.6 as: $\tau_s^t[\epsilon(r_s^t(e_s^t)) - \epsilon(r_{s'}^t(e_{s'}^t))] \geq \tau_{s'}^t[\epsilon(r_s^t(e_s^t)) - \epsilon(r_{s'}^t(e_{s'}^t))]$ and we conclude that $\tau_s^t > \tau_{s'}^t$. Similarly, we can prove $r_s^t = r_{s'}^t \Leftrightarrow \tau_s^t = \tau_{s'}^t$. $\square$

Proposition 2. *(Monotonicity) An optimal contract should have monotonic behavior, i.e., a seller of a higher type will sell more energy and receive a higher reward.*

Proof. A seller of higher type receives a higher reward based on Proposition 1, i.e., $r_1^t < \cdots < r_s^t < \cdots < r_{|S|}^t \Leftrightarrow \tau_1^t < \cdots < \tau_s^t < \cdots < \tau_{|S|}^t$. Then, based on the monotonic relationship among the reward r_s^t and the effort e_s^t, i.e., $r_s^t = \tau_s^t e_s^t$, we conclude that $e_1^t < \cdots < e_s^t < \cdots < e_{|S|}^t$. $\square$

Proposition 3. *(Rationality) An optimal contract should be rational, i.e., a seller of higher type should enjoy a higher utility.*

Proof. We write the IC condition for two indicative sellers $s \neq s'$, $\forall s,\ s' \in S$: $\tau_s^t \epsilon(e_s^t) - p_{\mathcal{S}} e_s^t \geq \tau_s^t \epsilon(e_{s'}^t) - p_{\mathcal{S}} e_{s'}^t \xLeftrightarrow{\tau_s^t > \tau_{s'}^t} \tau_s^t \epsilon(e_s^t) - p_{\mathcal{S}} e_s^t \geq \tau_{s'}^t \epsilon(e_{s'}^t) - p_{\mathcal{S}} e_{s'}^t \Leftrightarrow U_s^t(e_s^t) \geq$

$U^t_{s'}(e^t_{s'})$. We generalize this outcome for any seller s, $\forall s \in S$: $\tau^t_1 < \cdots < \tau^t_s < \cdots < \tau^t_{|S|} \Leftrightarrow U^t_1 < \cdots < U^t_s < \cdots < U^t_{|S|}$. $\square$

The physical meaning of the latter three conditions, i.e., fairness, monotonicity, and rationality, is that an optimal contract (e^{t*}_s, r^{t*}_s) should guarantee all of them in order to incentivize the sellers to sell part or all of their energy surplus during each time slot t, instead of locally storing it for future use. Based on the above analysis, the interactions among the MGO and the sellers can be captured as a contract-theoretic optimization problem aimed at determining the optimal personalized contracts. The optimization problem aims at jointly maximizing the MGO's profit (Eq. 8.7a), while guaranteeing the IR (Eq. 8.7b), IC (Eq. 8.7c), and fairness, monotonicity, and rationality conditions (Eq. 8.7d), and it is formally stated as follows.

$$\max_{\{e^t_s, r^t_s\}_{\forall s \in S}} \sum_{s=1}^{|S|} [Pr^{(t)}_s (e^t_s - r_s(e^t_s))] \tag{8.7a}$$

$$\text{s.t.} \quad \tau^t_s \epsilon(e^t_s) - p_{\mathcal{S}} e^t_s \geq 0, \forall s \in S \quad \text{(IR)} \tag{8.7b}$$

$$\tau^t_s \epsilon(e^t_s) - p_{\mathcal{S}} e^t_s \geq \tau^t_s \epsilon(e^t_{s'}) - p_{\mathcal{S}} e^t_{s'}, \forall s \neq s', s, s' \in S \quad \text{(IC)} \tag{8.7c}$$

$$0 \leq r^t_1 < r^t_2 < \cdots < r^t_s < \cdots < r^t_{|S|} \tag{8.7d}$$

The above optimization is clearly non-convex. Thus, we will reduce its constraints and rewrite it as a convex optimization problem to allow for a tractable and feasible solution. Starting with the IR constraint (Eq. 8.7b) and based on the IC and monotonicity conditions, we have: $\tau^t_s \epsilon(e^t_s) - p_{\mathcal{S}} e^t_s \geq \tau^t_s \epsilon(e^t_1) - p_{\mathcal{S}} e^t_1$, $\forall s \in S$. Also, we know that $\tau^t_s > \tau^t_1$, $\forall s \in S$, thus, $\tau^t_s \epsilon(e^t_s) - p_{\mathcal{S}} e^t_s \geq \tau^t_1 \epsilon(e^t_1) - p_{\mathcal{S}} e^t_1 \geq 0$. Also, given that the MGO provides just-sufficient rewards to incentivize the sellers to participate in the microgrid, we can equivalently replace the constraint in Eq. 8.7b with $\tau^t_1 \epsilon(e^t_1) - p_{\mathcal{S}} e^t_1 = 0$. Focusing on the reduction of the IC constraints (Eq. 8.7c), we introduce the following terminology: (i) s, s', $s' \in \{1, \ldots, s-1\}$: downward

IC constraints, (ii) s, $s-1$, $\forall s \in S$: local downward IC constraints, (iii) s, s', $s' \in \{s+1, \dots, |S|\}$: upward IC constraints, and (iv) s, $s+1$, $\forall s \in S$: local upwards IC constraints.

Lemma 1. *All the downward IC constraints are captured by the local downward IC constraints.*

Proof. We write the IC conditions for three sellers, $s-1$, s, $s+1$, as follows: $\tau^t_{s+1}\epsilon(e^t_{s+1}) - p_{\mathcal{S}}e^t_{s+1} \geq \tau^t_{s+1}\epsilon(e^t_s) - p_{\mathcal{S}}e^t_s$ and $\tau^t_s\epsilon(e^t_s) - p_{\mathcal{S}}e^t_s \geq \tau^t_s\epsilon(e^t_{s-1}) - p_{\mathcal{S}}e^t_{s-1}$. We know that $e^t_s > e^t_{s-1} \overset{\epsilon\nearrow}{\Longleftrightarrow} \epsilon(e^t_s) > \epsilon(e^t_{s-1}) \overset{\tau^t_{s+1}>\tau^t_s}{\Longleftrightarrow} \tau^t_{s+1}[\epsilon(e^t_s) - \epsilon(e^t_{s-1})] > \tau^t_s[\epsilon(e^t_s) - \epsilon(e^t_{s-1})] \geq p_{\mathcal{S}}(e^t_s - e^t_{s-1})$. We apply recursively the latter outcome for all the sellers: $\tau^t_{s+1}\epsilon(e^t_{s+1}) - p_{\mathcal{S}}e^t_{s+1} \geq \tau^t_{s+1}\epsilon(e^t_{s-1}) - p_{\mathcal{S}}e^t_{s-1} \geq \cdots \geq \tau^t_{s+1}\epsilon(e^t_1) - p_{\mathcal{S}}e^t_1$. Thus, we conclude that $\tau^t_s\epsilon(e^t_s) - p_{\mathcal{S}}e^t_s \geq \tau^t_s\epsilon(e^t_{s-1}) - p_{\mathcal{S}}e^t_{s-1}$, i.e., all the downward IC constraints are captured by the local downward IC constraints. □

Lemma 2. *All the upward IC constraints are captured by the local downward IC constraint.*

Proof. We write again the IC conditions for three indicative sellers, $s-1$, s, $s+1$, as follows:

$$\tau^t_{s-1}\epsilon(e^t_{s-1}) - p_{\mathcal{S}}e^t_{s-1} \geq \tau^t_{s-1}\epsilon(e^t_s) - p_{\mathcal{S}}e^t_s \tag{8.8}$$

$$\tau^t_s\epsilon(e^t_s) - p_{\mathcal{S}}e^t_s \geq \tau^t_s\epsilon(e^t_{s+1}) - p_{\mathcal{S}}e^t_{s+1} \tag{8.9}$$

Based on Eq. 8.9 and the fairness condition, we have the following expression:

$$\begin{aligned} p^t_s(e^t_{s+1} - e^t_s) &\geq \tau^t_s[\epsilon(e^t_{s+1}) - \epsilon(e^t_s)] \\ &\geq^{\tau^t_s \geq \tau^t_{s-1}} \tau^t_{s-1}[\epsilon(e^t_{s+1}) - \epsilon(e^t_s)] \end{aligned} \tag{8.10}$$

Based on Eq. 8.8, 8.10, we have: $\tau^t_{s-1}\epsilon(e^t_{s-1}) - p_{\mathcal{S}}e^t_{s-1} \geq \tau^t_{s-1}\epsilon(e^t_s) - p_{\mathcal{S}}e^t_s \geq \tau^t_{s-1}\epsilon(e^t_{s+1}) - p_{\mathcal{S}}e^t_{s+1}$. Thus, $\tau^t_{s-1}\epsilon(e^t_{s-1}) - p_{\mathcal{S}}e^t_{s-1} \geq \tau^t_{s-1}\epsilon(e^t_{s+1}) - p_{\mathcal{S}}e^t_{s+1}$, showing that all the upward IC constraints hold true, if the IC condition is satisfied for the seller with type τ^t_{s-1}. We apply recursively this outcome: $\tau^t_{s-1}\epsilon(e^t_{s-1}) - p_{\mathcal{S}}e^t_{s-1} \geq \tau^t_{s-1}\epsilon(e^t_{s+1}) - p_{\mathcal{S}}e^t_{s+1} \geq \cdots \geq \tau^t_{s-1}\epsilon(e^t_{|S|}) - p_{\mathcal{S}}e^t_{|S|}$. Thus, all the upward IC constraints are captured by the local downward IC constraints. □

Based on the above analysis of the reduction of the IR and IC constraints, we can rewrite the contract-theoretic optimization problem (8.7a)-(8.7d) as follows:

$$\max_{\{e^t_s, r^t_s\}_{\forall s \in S}} \sum_{s=1}^{|S|} [Pr^{(t)}_s(e^t_s - r_s(e^t_s))] \tag{8.11a}$$

$$\text{s.t.} \quad \tau^t_1\epsilon(e^t_1) - p_{\mathcal{S}}e^t_1 \geq 0 \tag{8.11b}$$

$$\tau^t_s\epsilon(e^t_s) - p_{\mathcal{S}}e^t_s = \tau^t_s\epsilon(e^t_{s-1}) - p_{\mathcal{S}}e^t_{s-1} \tag{8.11c}$$

$$0 \leq r^t_1 < r^t_2 < \cdots < r^t_s < \cdots < r^t_{|S|} \tag{8.11d}$$

The optimization problem (8.11a)-(8.11d) is a convex optimization problem and the optimal contract (e^{t*}_s, r^{t*}_s) can be determined based on standard convex optimization methods [85]. Detailed numerical results are presented in Section 8.5.

8.4 Buyers' & Electricity Market's Interactions

In this section, we focus on capturing the interactions of the buyers with the MGO. The goal of each buyer $\beta, \forall\beta \in B \subseteq N$ is to purchase the remaining amount of energy $(d^t_\beta - g^t_\beta - b^{t-1}_\beta)$ [kWh], that cannot be supported by her local energy generation. The MGO aims to incentivize the buyers to buy the total amount of energy that they need, by providing personalized rewards r^t_β. In our proposed approach, the interactions between the MGO and the buyers are captured via a contract-theoretic model.

In particular, the buyers invest an "effort" $e^t_\beta \in [0,1]$, which represents the percentage of energy that they buy with respect to their total energy need, i.e., $d^t_\beta - g^t_\beta - b^{t-1}_\beta$. Each buyer is characterized by a type $\tau^t_\beta = \frac{d^t_\beta - g^t_\beta - b^{t-1}_\beta}{\max_{\beta \in B}\{d^t_\beta - g^t_\beta - b^{t-1}_\beta\}}$, showing its relative potential compared to the rest of the buyers in terms of buying energy. The MGO offers personalized rewards $r^t_\beta = \tau^t_\beta e^t_\beta$ to each buyer, e.g., fixed energy price, in order to incentivize them to buy energy and not postpone or decrease their energy needs. The buyers' utility is defined as the gained profit from buying energy from the MGO and is defined as follows:

$$U^t_\beta(e^t_\beta) = \tau^t_\beta f(e^t_\beta) - p_M e^t_\beta. \tag{8.12}$$

The first term of Eq. 8.12 captures the buyer's personalized satisfaction from purchasing energy, where $f(e^t_\beta)$ is the buyer's satisfaction function, e.g., $f(e^t_\beta) = \sqrt{r^t_\beta(e^t_\beta)}$. The latter one captures the buyer's satisfaction from the consumption of the energy that they buy from the MGO. The buyer's satisfaction is a strictly increasing, continuous, and concave function with respect to the received reward r^t_β, as the buyer becomes more satisfied by covering more of their appliances' energy needs while such satisfaction becomes saturated at a specific upper limit of energy need. Also, $p_M \in [0,1]$ here is considered as a normalized dimensionless parameter representing the energy price, however in a real-life implementation it can be mapped to realistic values and units [$/kWh] [171]. The MGO's utility from selling energy to the buyers is obtained as its total profit, and id defined as follows:

$$U^t_{MGO,sell}(\mathbf{e_{buy}}) = \sum_{\beta=1}^{|B|} Pr^t_\beta [p_M e^t_\beta - r^t_\beta(e^t_\beta)] \tag{8.13}$$

As mentioned before, in the general case, the MGO has partial available information about the potential of each buyer to buy energy, thus, it probabilistically estimates each buyer's type τ_β^t with probability Pr_β^t, where $\sum_{\beta=1}^{|B|} Pr_\beta^t = 1$. Similarly to the sellers' case, several types of probability distributions can be adopted based on the nature of the examined energy market to realistically capture the buyers' characteristics. The goal of the MGO is to maximize its profit, while guaranteeing that the buyers will buy energy from the microgrid market. Considering the benchmarking scenario of complete information of the buyers' types, the interactions between the MGO and the buyers are formulated as a maximization problem of the MGO's profit (Eq. 8.14a), while considering the optimization of the buyers' utilities (Eq. 8.14b).

$$\max_{\{e_\beta^t\}_{\forall \beta \in B}} [p_M e_\beta^t - r_\beta^t(e_\beta^t)] \tag{8.14a}$$

$$\text{s.t.} \quad \tau_\beta^t f(e_\beta^t) - p_M e_\beta^t \geq 0, \forall \beta \in B \tag{8.14b}$$

Theorem 2. *The optimal contract among the MGO and each buyer β under the complete information scenario is* $(e_\beta^{t*}, r_\beta^{t*}) = (\frac{{\tau_\beta^t}^2}{2p_M}, \frac{{\tau_\beta^t}^2}{4})$.

Proof. It follows the same philosophy, reasoning and steps of Theorem 1. □

Under the realistic scenario of incomplete information regarding the buyers' types, the conditions of IR (Eq. 8.15b), IC (Eq. 8.15c), and fairness, monotonicity, and rationality (Eq. 8.15d) should hold true. Also, the optimal contract jointly maximizes the MGO's utility, i.e., profit, as follows.

$$\max_{\{e_\beta^t, r_\beta^t\}_{\forall \beta \in B}} \sum_{\beta=1}^{|B|} Pr_\beta^{(t)} [p_M e_\beta^t - r_\beta^t(e_\beta^t)] \tag{8.15a}$$

$$\text{s.t.} \quad \tau_\beta^t f(e_\beta^t) - p_M e_\beta^t \geq 0, \forall \beta \in B \quad \text{(IR)} \tag{8.15b}$$

$$\tau_\beta^t f(e_\beta^t) - p_M e_\beta^t \geq \tau_\beta^t f(e_{\beta'}^t) - p_M e_{\beta'}^t, \forall \beta \neq \beta' \quad \text{(IC)} \tag{8.15c}$$

$$0 \leq r_1^t < r_2^t < \cdots < r_\beta^t < \cdots < r_{|B|}^t \tag{8.15d}$$

Following similar reasoning as in Section 8.3, we can rewrite the above optimization problem and solve it with standard convex optimization methods as follows:

$$\max_{\{e_\beta^t, r_\beta^t\}_{\forall \beta \in B}} \sum_{\beta=1}^{|B|} Pr_\beta^{(t)} [p_M e_\beta^t - r_\beta^t(e_\beta^t)] \tag{8.16a}$$

$$\text{s.t.} \quad \tau_1^t f(e_1^t) - p_M e_1^t = 0 \tag{8.16b}$$

$$\tau_\beta^t f(e_\beta^t) - p_M e_\beta^t \geq \tau_\beta^t f(e_{\beta-1}^t) - p_M e_{\beta-1}^t \tag{8.16c}$$

$$0 \leq r_1^t < r_2^t < \cdots < r_\beta^t < \cdots < r_{|B|}^t \tag{8.16d}$$

The solution of the above problem concludes to the optimal contracts $(e_\beta^{t*}, r_\beta^{t*}), \forall \beta \in B$, determining the amount of purchased energy of the buyers and the MGO's offered personalized rewards to the buyers.

8.5 Numerical Results

In this section, a detailed evaluation analysis of the proposed contract-theoretic approaches is presented, via modeling and simulation, in order to demonstrate and assess the sellers and buyers interactions with the microgrid operator. Specifically, the pure operation characteristics and performance of the proposed framework for both the sellers and the buyers are presented in Section 8.5.1. The behavior of the prosumers, in terms of acting either as sellers or buyers, is studied in more detail in Section 8.5.2 with respect to the energy price, the energy generation cost, and the prosumers' energy generation characteristics during a day. Finally, the joint behavior of the prosumers and the MGO throughout the day for different energy generation use case scenarios is studied in Section 8.5.3, towards demonstrating and gaining more insights about their tight interconnection and interactions.

In the rest of the simulation results, we consider the following parameters: $d_n^t \in [0.50, 1.50]$ kWh, $g_n^t \in [0, 2]$ kWh, $p_M = 2$, $p_{\mathcal{S}} = 2$ [172]. Furthermore, for demonstration purposes and unless otherwise explicitly stated, we examine the system operation for $|T| = 24$ hours and $|N| = 100$ prosumers. The probabilities Pr_s^t and Pr_β^t are obtained assuming that the sellers' and buyers' types follow uniform distributions. For all the presented numerical results, a Monte Carlo analysis has been performed of $10,000$ executions to receive more representative outcomes.

8.5.1 Pure Operation Performance

Initially, the pure operation performance of the proposed prosumer-centric self-sustained smart grid system model based on the contract-theoretic approach is presented to capture the interactions of both the sellers (Section 8.3) and the buyers (Section 8.4) with the MGO. One indicative time slot t is considered, where both types of interactions are analyzed.

In particular, Fig. 8.2a - 8.2d present the sellers' and buyers' types (τ_s^t, τ_β^t), efforts (e_s^t, e_β^t), rewards (r_s^t, r_β^t), and utilities (U_s^t, U_β^t), respectively, under the scenarios of complete and incomplete information of the prosumers' types from the MGO's perspective. For demonstration purposes, the sellers' and buyers' IDs have been sorted with respect to their increasing types. The results show that the higher the seller's type is (Fig. 8.2a), the more energy surplus it has, thus, it is incentivized more by the MGO to sell its available energy by being offered a higher reward (Fig. 8.2c - left vertical axis). Consequently, it appears that indeed it sells more energy by investing a greater effort (Fig. 8.2b - left vertical axis). Thus, the seller of greater

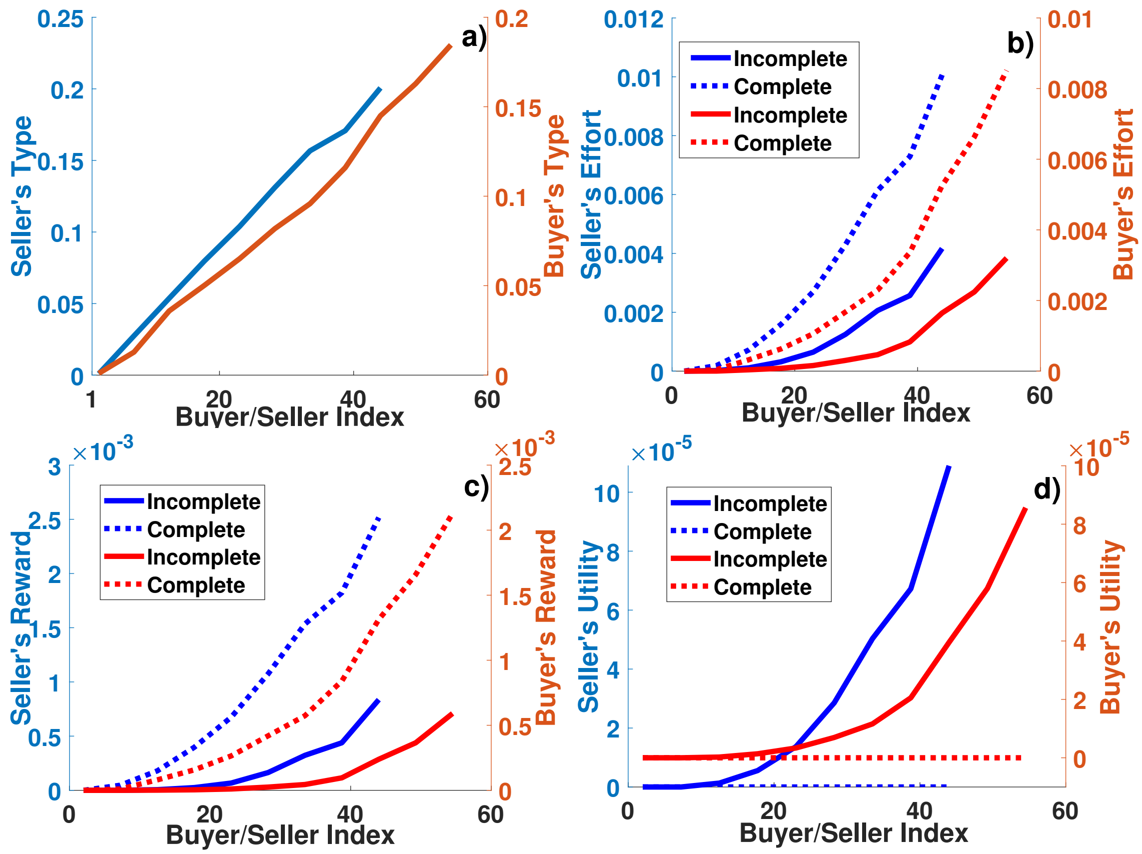

Figure 8.2: Sellers' and buyers' types, efforts, rewards, and utilities with respect to their index under the complete and incomplete information scenarios.

energy surplus ultimately achieves a higher utility (Fig. 8.2d - left vertical axis). With reference to the sellers, and by comparing the complete (i.e., benchmarking) and the incomplete (i.e., realistic) information scenarios, we observe that under the former, the MGO can fully exploit the sellers' energy surplus. This in turn means that the MGO provides to the sellers higher rewards to incentivize them to sell the vast majority of their available energy (Fig. 8.2c - left vertical axis), which indeed translates to having the sellers actually selling a higher amount of energy (Fig. 8.2b - left vertical axis), as compared to the incomplete information scenario.

Focusing our analysis on the buyers perspective and interactions with the MGO,

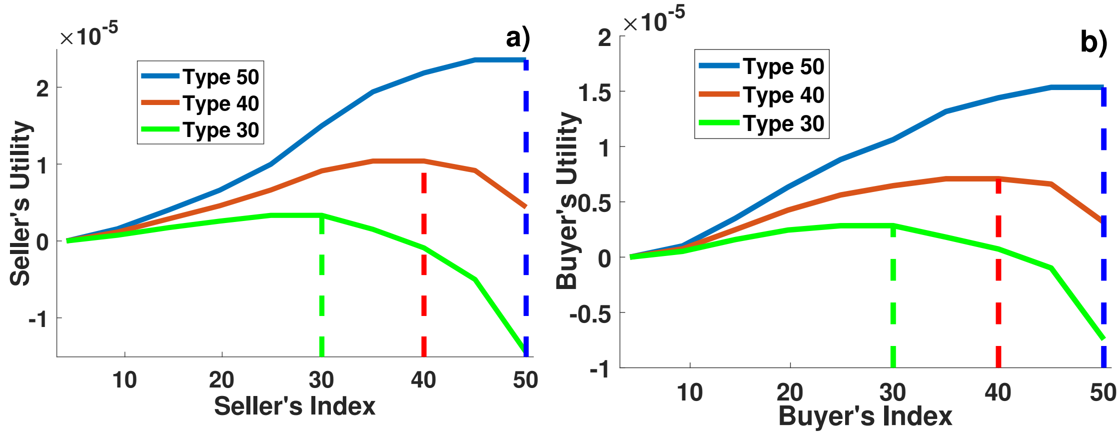

Figure 8.3: Incentive compatibility condition.

we observe that a buyer with higher need to purchase energy, i.e., of higher type (Fig. 8.2a), is incentivized more by the MGO to do so (Fig. 8.2c - right vertical axis). Thus, a buyer of higher type by ultimately purchasing more energy (Fig. 8.2b - right vertical axis), it covers the majority of its energy needs and accordingly achieves a higher utility (Fig. 8.2d - right vertical axis). Comparing the complete and incomplete information scenarios with reference to the buyers, the results confirm our theoretical analysis and observation, by clearly demonstrating that higher rewards are provided to the buyers (Fig. 8.2c - right vertical axis) in the complete information scenario, who purchase more energy (Fig. 8.2b - right vertical axis) compared to the incomplete information scenario. It should be highlighted that under the complete information scenario, both the sellers and the buyers achieve zero utility (Fig. 8.2d - right vertical axis), as the MGO provides just the sufficient rewards to marginally incentivize them to contribute to the microgrid's smooth and seamless operation.

Also, the results confirm that the individual rationality, incentive compatibility, fairness, monotonicity, and rationality conditions hold true for both the sellers and the buyers under all the examined scenarios. Specifically, Fig. 8.3a-8.3b present the sellers' and buyers' utility for the corresponding optimal contracts derived for each type for three indicative sellers and buyers with IDs $s = \{30, 40, 50\}$, and $\beta = \{30, 40, 50\}$, respectively. The results reveal that both the sellers and the buyers achieve the highest possible utility, when receiving the optimal personalized contract that is designed accounting for their unique energy generation, demand, and storage characteristics. This observation confirms the validity of the incentive compatibility condition. Also, the results show that a seller or buyer of a higher type achieves a higher utility, confirming the validity of the rationality condition.

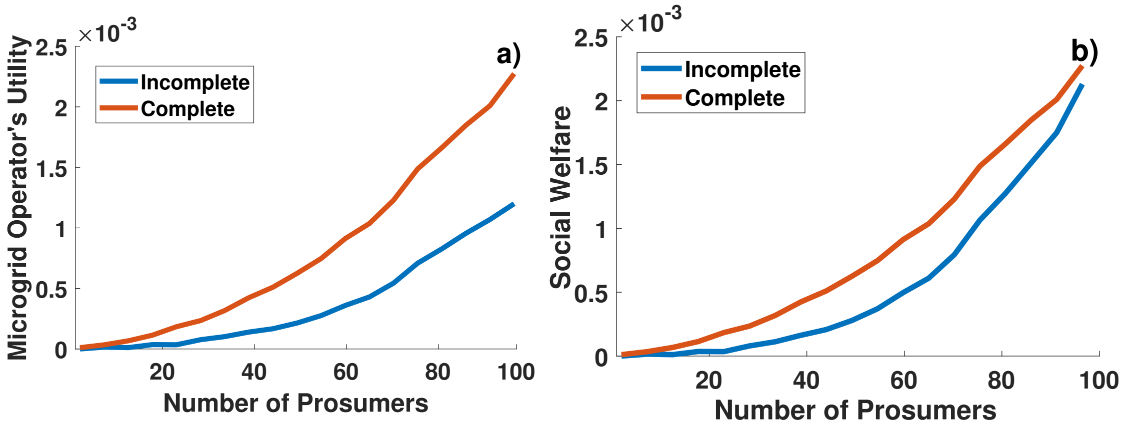

Figure 8.4: MGO's utility and social welfare under the complete and incomplete information scenarios.

Moreover, Fig. 8.4a-8.4b illustrate the MGO's utility and overall microgrid system's social welfare as a function of the number of prosumers residing in the microgrid system under the complete and incomplete information scenarios. The results show that better MGO's utility and social welfare is achieved under the complete information scenario, as the MGO can provide more targeted rewards by knowing the sellers' and buyers' exact types. However, it is highlighted that the incomplete information scenario, which is a realistic implementation of the microgrid system, achieves acceptable social welfare, especially, for increasing number of prosumers, with only 7% worse social welfare compared to the complete information scenario, for the case of 100 prosumers.

8.5.2 Prosumer's Behavior throughout the Day

In this subsection, we study the impact of various system and prosumer characteristics, such as energy price p_M, energy generation cost $p_{\mathcal{S}}$, and prosumer's energy generation $g_n^t, \forall n \in N, \forall t \in T$, on the behavior of the prosumers in terms of acting as sellers or buyers, throughout the operation of a day.

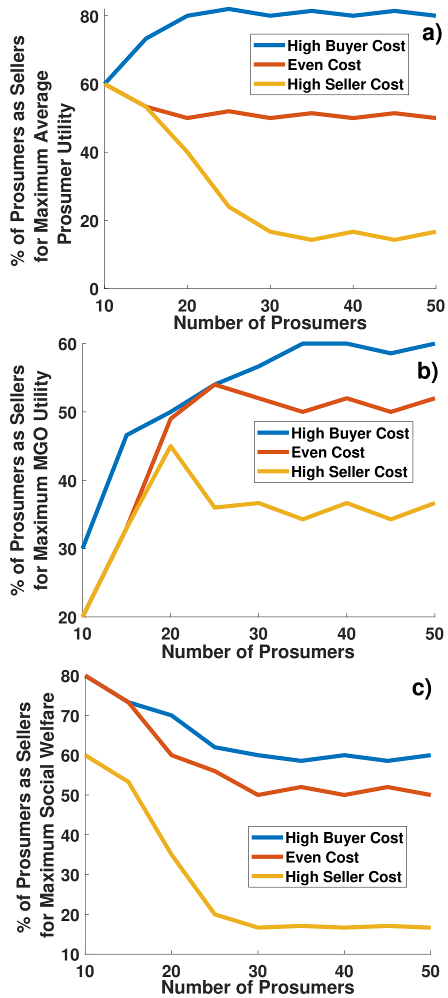

Figure 8.5: Percentage of prosumers acting as sellers to achieve maximum (a) prosumers' average utility (i.e., prosumer-centric), (b) MGO utility (i.e., MGO-centric), and (c) social welfare, vs. increasing number of prosumers, under three scenarios

In particular, Fig. 8.5a-8.5c present the percentage of prosumers that act as sellers in order to maximize the average prosumers' utility, the MGO's utility, and the social welfare, respectively, as a function of the number of prosumers in the microgrid. Three different scenarios of energy price p_M and energy generation costs $p_\mathcal{S}$ are considered: (i) Even Cost with $p_M = p_\mathcal{S} = 2$, (ii) High Buyer Cost with $p_M = 6, p_\mathcal{S} = 2$, and (iii) High Seller Cost with $p_M = 2, p_\mathcal{S} = 6$. The results reveal that when the energy generation cost is high (High Seller Cost scenario), a smaller percentage of sellers is incentivized to sell energy, as their energy production cost is high, and the sellers prefer to keep their generated energy for future use. In contrast, when the energy price is high (High Buyer Cost scenario), the sellers can achieve a higher profit by selling their generated energy to the MGO, thus, a higher percentage of prosumers acts as sellers. The Even Cost scenario presents an intermediate behavior between the High Seller and the High Buyer Cost scenarios. Also, comparing the prosumer-centric approach (Fig. 8.5a) against an MGO-centric approach (Fig. 8.5b) that aims at maximizing the MGO utility, we observe in the latter case a higher offset of the percentages of the prosumers acting as sellers for the High Seller and Even Cost scenarios, as the MGO aggressively provides rewards to the prosumers to sell their energy. The opposite holds true for the High Buyer scenario, as the MGO prefers to sell its available energy to the buyers at a higher price, as compared to buying energy from the sellers. An intermediate behavior of the percentages of the prosumers acting as sellers is observed when the goal is to solely maximize their social welfare (Fig. 8.5c), as the selfish behavior of the MGO and the prosumers is balanced.

Fig. 8.6a-8.6b present the prosumers' average energy generation and the percentage of them that act as sellers during the day, respectively, for two comparative scenarios: (i) High Generation scenario, and (ii) Low Generation scenario, where the prosumers have high and low energy generation capacity, respectively. The prosumers' energy generation is solely based on solar photovoltaic panels. The results reveal that during the sunny periods of the day, the prosumers generate more energy

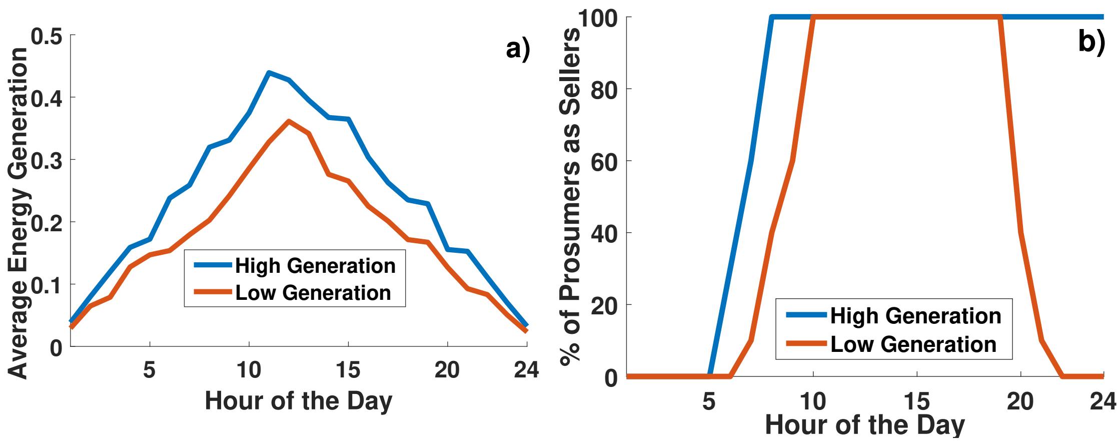

Figure 8.6: Average energy generation and percentage of prosumers acting as sellers under high and low energy generation

(Fig. 8.6a), thus, a greater percentage of them is incentivized to act as sellers (Fig. 8.6b) in both examined scenarios. Additionally, in the High Generation scenario, it is observed that the prosumers generate sufficient amount of energy to cover their personal energy needs, thus, they act as sellers for the majority of the day's duration. It is noted that the energy price and the energy generation costs are assumed to remain fixed throughout the day.

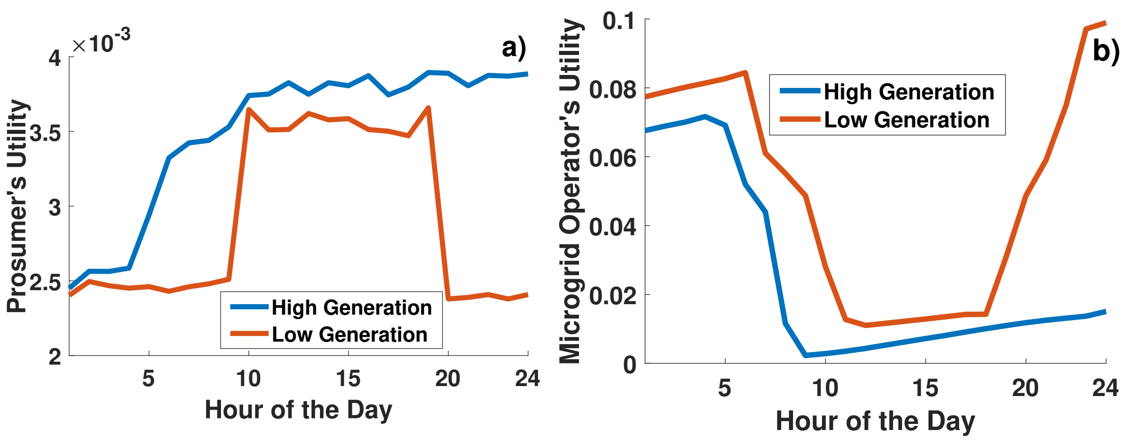

Figure 8.7: Average prosumers' utility and MGO's utility for $p_M = 6, p_S = 2$ under high and low energy generation

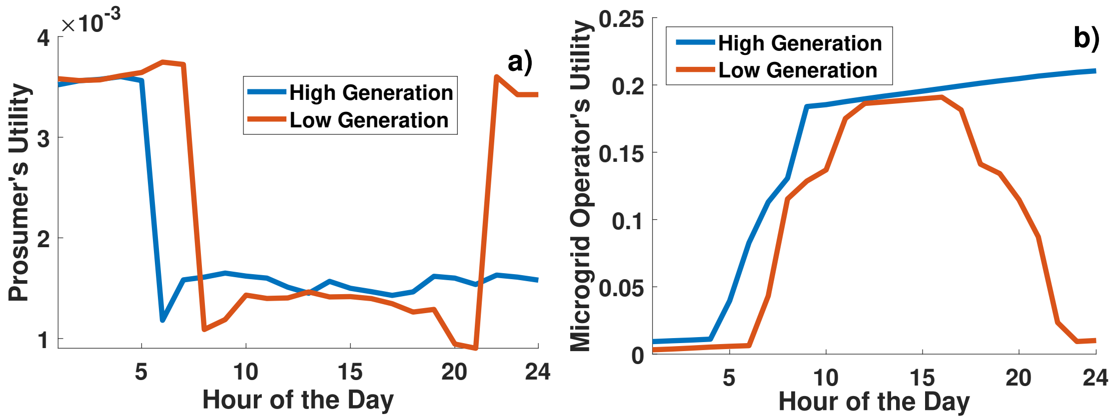

Figure 8.8: Average prosumers' utility and MGO's utility for $p_M = 2, p_S = 6$ under high and low energy generation scenarios

The above analysis and evaluation is further extended in Fig. 8.7 and Fig. 8.8 where both the prosumers and MGO average utilities are presented for two different scenarios: High Buyer Cost with $p_M = 6, p_S = 2$ (Fig. 8.7), and High Seller Cost with $p_M = 2, p_S = 6$ (Fig. 8.8), respectively. The energy generation characteristics of the prosumers for both scenarios follow the behavior presented in Fig. 8.6a. The results reveal that when the energy price p_M is high and the energy generation cost p_S is low, more prosumers act as sellers, thus, their average utility is higher (Fig. 8.7a), compared to the alternative scenario (Fig. 8.8a). This trend is expected as the prosumers generate energy with low cost. The offset of the High and Low Generation scenarios in Fig. 8.7a and Fig. 8.8a stems from the corresponding percentage of prosumers that act as sellers. Also, the benefit of the prosumers corresponds to the loss of the MGO, thus, the exact flipped trend is observed in Fig. 8.7b and Fig. 8.8b regarding the achieved MGO's utility.

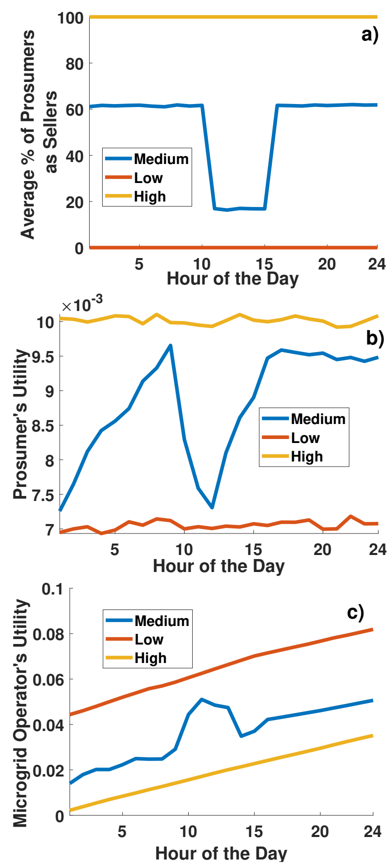

Figure 8.9: Percentage of prosumers acting as sellers, prosumers' average utility, and MGO's utility during the day for the three energy generation scenarios ($p_M = 6, p_{\mathcal{S}} = 2$)

8.5.3 Impact of Energy Generation and Demand on the Prosumers' and MGO's Interactions

In the following, we focus our study on the impact of the prosumers' energy generation and demand characteristics on their interactions with the microgrid system during a day. We consider an evolving behavior where the prosumers' demand is low for $t \in [0, 8]$, then it increases for $t \in [8, 16]$, and then drops again for $t \in [16, 24]$, representing the realistic prosumers' energy demand during the day. The aforementioned case, is evaluated and studied under three different scenarios regarding the prosumers' energy generation capacity, i.e, High, Medium, and Low. Following a similar methodology with our evaluation in the previous figures in this subsection (i.e., considering high buyer cost and high seller cost alternatives), two sets of results are produced and presented, differentiated exactly with respect to the considered energy price p_M and the energy generation cost $p_{\mathcal{S}}$, i.e., $p_M = 6, p_{\mathcal{S}} = 2$ for Fig. 8.9-8.10 and $p_M = 2, p_{\mathcal{S}} = 6$ for Fig. 8.11-8.12.

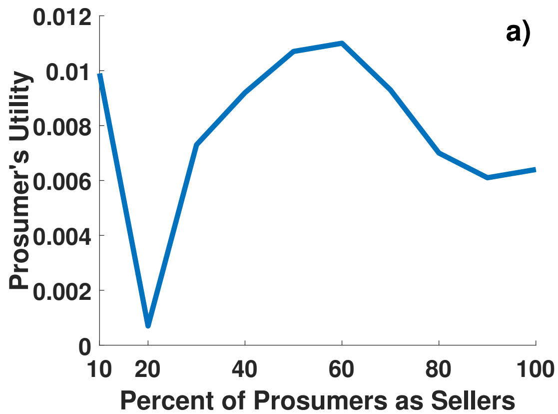

Figure 8.10: Achieved utility in the system of Fig. 8.9 for different percentages of prosumers functioning as sellers.

Specifically, Fig. 8.9a-8.9c present the percentage of prosumers acting as sellers, their average utility and the MGO's utility during the day, respectively, for the three aforementioned energy generation scenarios. The results reveal that in the High energy generation scenario, all the prosumers act as sellers, as they have sufficient energy surplus to support their personal energy need, while the exact opposite holds true in the Low energy generation scenario (Fig. 8.9a). In order to gain insight about the behavior of the curves for the Medium energy generation scenario, we should turn our attention to Fig. 8.10.

Particularly, with reference to the Medium energy generation scenario, Fig. 8.10 presents the behavior of the prosumers' average utility (Fig. 8.10a) and the MGO's utility (Fig. 8.10b) for different percentages of prosumers acting as sellers (horizontal axis). We observe that the maximum prosumers' average utility is achieved, when approximately 60% of the prosumers act as sellers. Correlating this value with the results in Fig. 8.9a, we observe that approximately 60% of the prosumers act as sellers, except for the time interval 11 am - 3 pm, when the prosumers' energy demand becomes high, and accordingly fewer prosumers are acting as sellers. Also, given that the energy generation cost is low, i.e., $p_S = 2$, as compared to the energy price, which is high, i.e., $p_M = 6$, the prosumers that generate a lot of energy achieve higher utility compared to the scenario of generating small amount of energy (Fig. 8.9b). The exact opposite is observed from the MGO's perspective (Fig. 8.9c). Focusing on the Medium energy generation scenario, we observe that during the morning hours, i.e., $t \in [0, 11]$, the prosumers generate more energy, thus accumulating energy surplus, and approximately 60% of them act as sellers (Fig. 8.9a), achieving the maximum possible utility (Fig. 8.10a), thus, their average achieved utility increases during this time frame (Fig. 8.9b). In the slot $t \in [10, 11]$, their energy demands increase and the percentage of prosumers acting as sellers drops to 20% (Fig. 8.9a), achieving the lowest possible utility (Fig. 8.10a), driving their average utility to drop (Fig. 8.9b). Similar analysis and reasoning can be derived and followed for the rest of the day.

Still focusing on the Medium energy generation scenario but from the MGO's perspective, we jointly study Fig. 8.9c and Fig. 8.10b. We observe that for the time periods $t \in [0, 10] \cup [15, 24]$, where the percentage of the prosumers acting as sellers is approximately 60% (Fig. 8.9a), the MGO achieves a relatively high utility (Fig. 8.10b), and slowly increases its profit during those time periods. On the other hand, for the time interval $t \in [10, 15]$, where only 20% of the prosumers act as sellers (Fig. 8.9a) due to their personal high energy demand, the MGO achieves the highest possible utility (Fig. 8.10b), as the MGO has set its price at a high value, i.e., $p_M = 6$. Thus, during this period, the MGO accumulates a higher profit.

Last, a symmetric scenario is presented in Fig. 8.11 and Fig. 8.12, where a

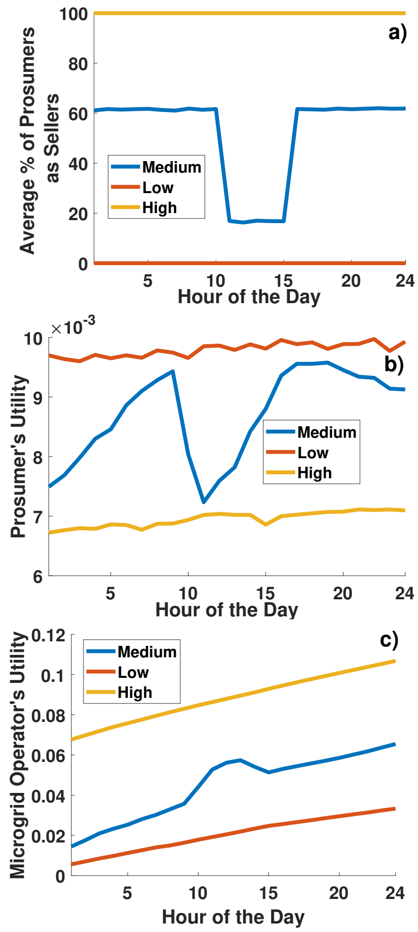

Figure 8.11: Percentage of prosumers acting as sellers, prosumers' average utility, and MGO's utility during the day for the three energy generation scenarios ($p_M = 2, p_S = 6$)

low energy price ($p_M = 2$) and high energy generation cost ($p_S = 6$) is considered instead. The point that should be highlight here is that the average utility of the prosumers with high energy generation is lower compared to the ones with low energy generation as the energy generation cost is higher. The exact opposite behavior is presented by the MGO's utility, as more prosumers tend to buy energy from it.

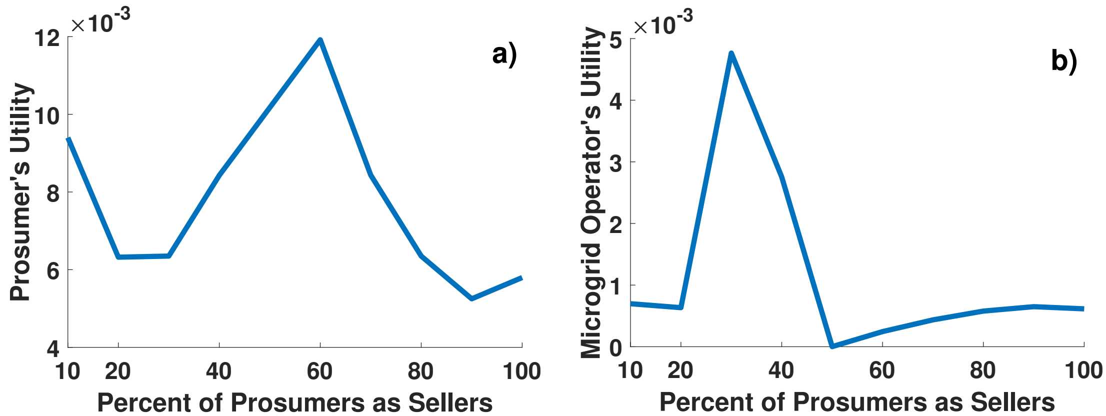

Figure 8.12: Achieved utility in the system of Fig. 8.11 for different percentages of prosumers functioning as sellers.

8.6 Conclusions

In this chapter, the paradigm of prosumer-centric self-sustained smart grid systems is introduced, by capturing and properly modeling the interactions of the prosumers with the microgrid operator via a labor economics based approach. The prosumers throughout the system operation may serve as sellers or buyers, based on their personal energy generation, demand, and storage characteristics. The MGO offers personalized rewards to the sellers and buyers to incentivize them to sell and purchase energy, respectively. The contract-theoretic optimization problems between the MGO and the sellers and the MGO and the buyers respectively, are formulated and solved to determine the optimal personalized contracts, i.e., rewards and amount of sold/purchased energy. Detailed numerical and comparative evaluation results - obtained via modeling and simulation - are presented to demonstrate the operation of the proposed framework and highlight its main characteristics, under various diverse scenarios.

Part of our current and future work is to incorporate the prosumers reactions to the energy price fluctuation, which introduces risk in their decision to act as sellers or buyers. In our efforts to study and address this problem, the principles of Prospect Theory are adopted. Furthermore, along the same lines, we plan to consider a system where multiple MGOs may co-exist and therefore the prosumers may dynamically get associated with different MGOs at different times, thus introducing several uncertainties within a more competitive market overall.

Chapter 9

Conclusion and Future Works

This dissertation proposes novel theory and scalable algorithms for decentralized intelligent decision making models, frameworks, and algorithms to support the smooth operation of Cyber Physical Social Systems.

9.0.1 Summary of Contributions

In this dissertation, we focus our research activities on devising decentralized intelligent decision making models, frameworks, and algorithms to support the smooth operation of Cyber Physical Social Systems. The proposed decentralized intelligent decision making models are jointly exploiting theories from the field of Economics, such as Game Theory and Contract Theory, and from the field of Computer Science, such as Reinforcement Learning concepts. Reinforcement learning is applied to allow for humans to make informed decisions in the considered Cyber Physical Social Systems based off of the environment around them. Additionally, contract theoretic and game theoretic models allow for us to accurately depict the relationships between the different involved entities in the examined system.

Five main research problems have been examined in this Ph.D. dissertation, which can be summarized as follows:

- socio-physical human orchestration in smart cities,
- socio-aware public safety framework design,
- unmanned aerial vehicle or UAV-enabled dynamic multi-target tracking and sensing framework,
- resource orchestration in wireless powered communication public safety systems, and
- health data acquisition from wearable devices during a pandemic by following a techno-economics approach
- museum and visitors interactions enabled by labor economics
- prosumer-centric self-sustained smart grid systems

9.0.2 Future Work

The work summarized in this Ph.D. dissertation proposes a meaningful and general framework, where the control intelligence and decision making process is done by the humans towards sophisticated sensing the dynamic environment and making autonomous decisions. Part of our current and future work targets at addressing the problem of modeling and orchestrating the interactions in a variety of different cyber-physical-social systems. As each CPSS is unique each one may need to have the model utilized tailored to that specific CPSS. For example in Chapter 7, the museum operator provides monetary rewards to the visitors in exchange for their contributions, which are expressed as their total number of provided feedback evaluations of visited exhibits over their touring time. The interactions among the museum operator and visitors are captured in appropriately designed utility functions following the principles of labor economics, while the visitors' behavioral characteristics are utilized to define their unique types. Under such a setting and formulation, the goal of the museum operator is to optimize their profit and benefits, while jointly satisfying the visitors' quality of experience prerequisites, as reflected via their utility functions. The corresponding optimization problem is treated and solved under the general and realistic case of incomplete information, wherein the museum operator estimates the visitors' types probabilistically. The resulting outcome, referred to as "optimal contract" jointly determines the visitors' optimal contributions, as well as the museum operator's optimal amount of personalized rewards provided to each visitor. This model follows the general trend that has been utilized for other CPSS but required some slight adjustments to fit that specific CPSS, such as understanding the different ways museum visitors view and evaluate the museum. These adjustments allowed us to create a more holistic approach. Applying the general approach and tuning the approach to more CPSSs, as mentioned previously, is the basis of our future work.

References

[1] E. A. Lee, "Cyber physical systems: Design challenges," in *2008 11th IEEE international symposium on object and component-oriented real-time distributed computing (ISORC)*, pp. 363–369, IEEE, 2008.

[2] L. D. Xu and L. Duan, "Big data for cyber physical systems in industry 4.0: a survey," *Enterprise Information Systems*, vol. 13, no. 2, pp. 148–169, 2019.

[3] H. Habibzadeh, B. H. Nussbaum, F. Anjomshoa, B. Kantarci, and T. Soyata, "A survey on cybersecurity, data privacy, and policy issues in cyber-physical system deployments in smart cities," *Sustainable Cities and Society*, vol. 50, p. 101660, 2019.

[4] G. Schirner, D. Erdogmus, K. Chowdhury, and T. Padir, "The future of human-in-the-loop cyber-physical systems," *Computer*, vol. 46, no. 1, pp. 36–45, 2013.

[5] M. Gil, M. Albert, J. Fons, and V. Pelechano, "Designing human-in-the-loop autonomous cyber-physical systems," *International Journal of Human-Computer Studies*, vol. 130, pp. 21–39, 2019.

[6] J. Zeng, L. T. Yang, M. Lin, H. Ning, and J. Ma, "A survey: Cyber-physical-social systems and their system-level design methodology," *Future Generation Computer Systems*, vol. 105, pp. 1028–1042, 2020.

[7] J. J. Zhang, F.-Y. Wang, X. Wang, G. Xiong, F. Zhu, Y. Lv, J. Hou, S. Han, Y. Yuan, Q. Lu, *et al.*, "Cyber-physical-social systems: The state of the art and perspectives," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 3, pp. 829–840, 2018.

[8] Y. Zhou, F. R. Yu, J. Chen, and Y. Kuo, "Cyber-physical-social systems: A state-of-the-art survey, challenges and opportunities," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 1, pp. 389–425, 2019.

[9] A. Smirnov, A. Kashevnik, and A. Ponomarev, "Multi-level self-organization in cyber-physical-social systems: Smart home cleaning scenario," *Procedia Cirp*, vol. 30, pp. 329–334, 2015.

[10] E. E. Tsiropoulou, J. S. Baras, S. Papavassiliou, and S. Sinha, "Rfid-based smart parking management system," *Cyber-Physical Systems*, vol. 3, no. 1-4, pp. 22–41, 2017.

[11] S. Stoyanov, D. Orozova, and I. Popchev, "Internet of things water monitoring for a smart seaside city," in *2018 20th International Symposium on Electrical Apparatus and Technologies (SIELA)*, pp. 1–3, 2018.

[12] N. Yamanaka, G. Yamamoto, S. Okamoto, and A. Fumagalli, "Cyber-physical system for autonomous driving vehicle considering with social welfare," in *2020 22nd International Conference on Transparent Optical Networks (ICTON)*, pp. 1–4, 2020.

[13] 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, no. 1, pp. 283–326, 2018.

[14] S. Kang, C. Jeong, and K. Chung, "Tree-based real-time advertisement recommendation system in online broadcasting," *IEEE Access*, vol. 8, pp. 192693–192702, 2020.

[15] 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, IEEE, 2015.

[16] G. Song, Y. Wen, Y. Jia, and H. Liu, "Research on medical service system based on big data technology," in *2019 International Conference on Intelligent Transportation, Big Data Smart City (ICITBS)*, pp. 302–304, 2019.

[17] S. Lu, A. Wang, S. Jing, T. Shan, X. Zhang, Y. Guo, and Y. Liu, "A study on service-oriented smart medical systems combined with key algorithms in the iot environment," *China Communications*, vol. 16, no. 9, pp. 235–249, 2019.

[18] P. A. Apostolopoulos, E. E. Tsiropoulou, and S. Papavassiliou, "Demand response management in smart grid networks: A two-stage game-theoretic learning-based approach," *Mobile Networks and Applications*, pp. 1–14, 2018.

[19] A. Ghosal and M. Conti, "Key management systems for smart grid advanced metering infrastructure: A survey," *IEEE Communications Surveys Tutorials*, vol. 21, no. 3, pp. 2831–2848, 2019.

[20] 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.

[21] M. R. M. Kassim, "Iot applications in smart agriculture: Issues and challenges," in *2020 IEEE Conference on Open Systems (ICOS)*, pp. 19–24, 2020.

[22] B. Citoni, F. Fioranelli, M. A. Imran, and Q. H. Abbasi, "Internet of things and lorawan-enabled future smart farming," *IEEE Internet of Things Magazine*, vol. 2, no. 4, pp. 14–19, 2019.

[23] M. Roopaei, P. Rad, and K. R. Choo, "Cloud of things in smart agriculture: Intelligent irrigation monitoring by thermal imaging," *IEEE Cloud Computing*, vol. 4, no. 1, pp. 10–15, 2017.

[24] E. Tsiropoulou, K. Koukas, and S. Papavassiliou, "A socio-physical and mobility-aware coalition formation mechanism in public safety networks," *EAI Endorsed Trans. Future Internet*, vol. 4, p. 154176, 2018.

[25] M. Mendonça, B. Moreira, J. Coelho, N. Cacho, F. Lopes, E. Cavalcante, A. Dias, J. L. Ribeiro, E. Loiola, D. Estaregue, and B. Moura, "Improving public safety at fingertips: A smart city experience," in *2016 IEEE International Smart Cities Conference (ISC2)*, pp. 1–6, 2016.

[26] 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, no. 1, p. 33, 2017.

[27] A. Chianese, F. Piccialli, and J. E. Jung, "The internet of cultural things: Towards a smart cultural heritage," in *2016 12th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*, pp. 493–496, 2016.

[28] H. P. McKenna, "Emergent ambient culture in smart cities: Exploring the internet of cultural things (ioct) and applications in 21st century urban spaces," in *2016 12th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*, pp. 420–427, 2016.

[29] E. E. Tsiropoulou, J. S. Baras, S. Papavassiliou, and G. Qu, "On the mitigation of interference imposed by intruders in passive rfid networks," in *International Conference on Decision and Game Theory for Security*, pp. 62–80, Springer, 2016.

[30] H. Yetis and M. Karakose, "A cyber-physical-social system based method for smart citizens in smart cities," in *2020 24th International Conference on Information Technology (IT)*, pp. 1–4, 2020.

[31] S. K. Sowe, E. Simmon, K. Zettsu, F. de Vaulx, and I. Bojanova, "Cyber-physical-human systems: Putting people in the loop," *IT Professional*, vol. 18, no. 1, pp. 10–13, 2016.

[32] E. E. Tsiropoulou, G. Mitsis, and S. Papavassiliou, "Interest-aware energy collection & resource management in machine to machine communications," *Ad Hoc Networks*, vol. 68, pp. 48–57, 2018.

[33] C. Singhal and S. De, *Resource allocation in next-generation broadband wireless access networks*. IGI Global, 2017.

[34] J. J. Zhang, F. Wang, X. Wang, G. Xiong, F. Zhu, Y. Lv, J. Hou, S. Han, Y. Yuan, Q. Lu, and Y. Lee, "Cyber-physical-social systems: The state of the art and perspectives," *IEEE Transactions on Computational Social Systems*, vol. 5, no. 3, pp. 829–840, 2018.

[35] M. T. Thai, W. Wu, and H. Xiong, *Big Data in Complex and Social Networks*. CRC Press, 2016.

[36] E. E. Tsiropoulou, S. T. Paruchuri, and J. S. Baras, "Interest, energy and physical-aware coalition formation and resource allocation in smart iot applications," in *2017 51st Annual Conference on Information Sciences and Systems (CISS)*, pp. 1–6, IEEE, 2017.

[37] D. Serpanos, "The cyber-physical systems revolution," *Computer*, vol. 51, no. 3, pp. 70–73, 2018.

[38] M. Gharib, P. Lollini, and A. Bondavalli, "Towards an approach for analyzing trust in cyber-physical-social systems," in *2017 12th System of Systems Engineering Conference (SoSE)*, pp. 1–6, 2017.

[39] P. Bolton, M. Dewatripont, *et al.*, *Contract theory*. MIT press, 2005.

[40] M. J. Osborne *et al.*, *An introduction to game theory*, vol. 3. Oxford university press New York, 2004.

[41] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.

[42] N. Patrizi, P. A. Apostolopoulos, K. Rael, and E. E. Tsiropoulou, "Socio-physical human orchestration in smart cities," in *2019 IEEE International Conference on Smart Computing (SMARTCOMP)*, pp. 115–120, IEEE, 2019.

[43] G. Fragkos, N. Patrizi, E. E. Tsiropoulou, and S. Papavassiliou, "Socio-aware public safety framework design: A contract theory based approach," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, pp. 1–7, IEEE, 2020.

[44] N. Patrizi, G. Fragkos, K. Ortiz, M. Oishi, and E. E. Tsiropoulou, "A uav-enabled dynamic multi-target tracking and sensing framework," in *GLOBECOM 2020-2020 IEEE Global Communications Conference*, pp. 1–6, IEEE, 2020.

[45] 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, IEEE, 2020.

[46] S. A. Smith, *Contract theory*. OUP Oxford, 2004.

[47] 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.

[48] M. Diamanti, E. E. Tsiropoulou, and S. Papavassiliou, "Resource orchestration in uav-assisted noma wireless networks: A labor economics perspective," in *ICC 2021-IEEE International Conference on Communications*, pp. 1–6, IEEE, 2021.

[49] 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, IEEE, 2021.

[50] Y. Zhang and M. Guizani, *Game theory for wireless communications and networking*. CRC press, 2011.

[51] T. Başar and G. J. Olsder, *Dynamic noncooperative game theory*. SIAM, 1998.

[52] E. E. Tsiropoulou, G. K. Katsinis, P. Vamvakas, and S. Papavassiliou, "Efficient uplink power control in multi-service two-tier femtocell networks via a game theoretic approach," in *2013 IEEE 18th International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, pp. 104–108, 2013.

[53] K. Ritzberger *et al.*, "Foundations of non-cooperative game theory," *OUP Catalogue*, 2002.

[54] Y. Kawamoto, N. Yamada, H. Nishiyama, N. Kato, Y. Shimizu, and Y. Zheng, "A feedback control-based crowd dynamics management in iot system," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1466–1476, 2017.

[55] S. Karanikolaou, I. Boutsis, and V. Kalogeraki, "Understanding event attendance through analysis of human crowd behavior in social networks," in *Proceedings of the 8th ACM International Conference on Distributed Event-Based Systems*, pp. 322–325, ACM, 2014.

[56] I. Boutsis, S. Karanikolaou, and V. Kalogeraki, "Personalized event recommendations using social networks," in *2015 16th IEEE International Conference on Mobile Data Management*, vol. 1, pp. 84–93, 2015.

[57] H. Wang, M. Terrovitis, and N. Mamoulis, "Location recommendation in location-based social networks using user check-in data," in *Proceedings of the 21st ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 374–383, ACM, 2013.

[58] L. Guo, J. Shao, K. L. Tan, and Y. Yang, "Wheretogo: Personalized travel recommendation for individuals and groups," in *IEEE 15th International Conference on Mobile Data Management*, vol. 1, pp. 49–58, 2014.

[59] J. She, Y. Tong, L. Chen, and T. Song, "Feedback-aware social event-participant arrangement," in *Proceedings of the 2017 ACM International Conference on Management of Data*, pp. 851–865, ACM, 2017.

[60] J. Huang, Y. Zhou, X. Jia, and H. Sun, "A novel social event organization approach for diverse user choices," *The Computer Journal*, vol. 60, no. 7, pp. 1078–1095, 2016.

[61] N. Bikakis, V. Kalogeraki, and D. Gunopulos, "Social event scheduling," in *IEEE 34th Intern. Conf. on Data Engineering*, pp. 1272–1275, 2018.

[62] T.-A. N. Pham, X. Li, G. Cong, and Z. Zhang, "A general graph-based model for recommendation in event-based social networks," in *IEEE 31st International Conference on Data Engineering*, pp. 567–578, 2015.

[63] A. S. Poznyak and K. Najim, *Learning automata and stochastic optimization*, vol. 3. Springer, 1997.

[64] E. E. Tsiropoulou, G. K. Katsinis, A. Filios, and S. Papavassiliou, "On the problem of optimal cell selection and uplink power control in open access multi-service two-tier femtocell networks," in *International Conference on Ad-Hoc Networks and Wireless*, pp. 114–127, Springer, 2014.

[65] R. Mirchandaney and J. A. Stankovic, "Using stochastic learning automata for job scheduling in distributed processing systems," *Journal of Parallel and Distributed Computing*, vol. 3, no. 4, pp. 527–552, 1986.

[66] P. A. Apostolopoulos, E. E. Tsiropoulou, and S. Papavassiliou, "Game-theoretic learning-based qos satisfaction in autonomous mobile edge computing," in *2018 Global Information Infrastructure and Networking Symposium (GIIS)*, pp. 1–5, IEEE, 2018.

[67] P. Vamvakas, E. E. Tsiropoulou, and S. Papavassiliou, "Risk-aware resource management in public safety networks," *Sensors*, vol. 19, no. 18, p. 3853, 2019.

[68] K. Starbird, J. Maddock, M. Orand, P. Achterman, and R. M. Mason, "Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing," *IConference 2014 Proceedings*, 2014.

[69] F. Bloch, G. Demange, and R. Kranton, "Rumors and social networks," *International Economic Review*, vol. 59, no. 2, pp. 421–448, 2018.

[70] D. E. Alexander, "Social media in disaster risk reduction and crisis management," *Science and engineering ethics*, vol. 20, no. 3, pp. 717–733, 2014.

[71] B. Resch, "People as sensors and collective sensing-contextual observations complementing geo-sensor network measurements," in *Progress in location-based services*, pp. 391–406, Springer, 2013.

[72] V. Krishnamurthy and H. V. Poor, "A tutorial on interactive sensing in social networks," *IEEE Transactions on Computational Social Systems*, vol. 1, no. 1, pp. 3–21, 2014.

[73] F. Amato, V. Moscato, A. Picariello, and G. Sperli'ì, "Extreme events management using multimedia social networks," *Future Generation Computer Systems*, vol. 94, pp. 444–452, 2019.

[74] X. Zhang, Z. Yang, Z. Zhou, H. Cai, L. Chen, and X. Li, "Free market of crowdsourcing: Incentive mechanism design for mobile sensing," *IEEE transactions on parallel and distributed systems*, vol. 25, no. 12, pp. 3190–3200, 2014.

[75] A. Goncalves, C. Silva, P. Morreale, and J. Bonafide, "Crowdsourcing for public safety," in *2014 IEEE International Systems Conference Proceedings*, pp. 50–56, IEEE, 2014.

[76] G. Fragkos, C. Minwalla, J. Plusquellic, and E. E. Tsiropoulou, "Artificially intelligent electronic money," *IEEE Consumer Electronics Magazine*, vol. 10, no. 4, pp. 81–89, 2020.

[77] M. Diamanti, G. Fragkos, E. E. Tsiropoulou, and S. Papavassiliou, "Resource orchestration in interference-limited small cell networks: A contract-theoretic approach," in *International Conference on Network Games, Control and Optimization*, pp. 101–109, Springer, 2021.

[78] F. Sangoleye, N. Irtija, and E. E. Tsiropoulou, "Smart energy harvesting for internet of things networks," *Sensors*, vol. 21, no. 8, p. 2755, 2021.

[79] Z. Hasan and V. K. Bhargava, "Relay selection for ofdm wireless systems under asymmetric information: A contract-theory based approach," *IEEE Transactions on Wireless Communications*, vol. 12, no. 8, pp. 3824–3837, 2013.

[80] Y. Zhang, L. Song, W. Saad, Z. Dawy, and Z. Han, "Contract-based incentive mechanisms for device-to-device communications in cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 10, pp. 2144–2155, 2015.

[81] L. Duan, L. Gao, and J. Huang, "Cooperative spectrum sharing: A contract-based approach," *IEEE Transactions on Mobile Computing*, vol. 13, no. 1, pp. 174–187, 2012.

[82] A. V. Kordali and P. G. Cottis, "A contract-based spectrum trading scheme for cognitive radio networks enabling hybrid access," *IEEE Access*, vol. 3, pp. 1531–1540, 2015.

[83] E. E. Tsiropoulou, P. Vamvakas, and S. Papavassiliou, "Energy efficient uplink joint resource allocation non-cooperative game with pricing," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2352–2356, IEEE, 2012.

[84] M. Doo and L. Liu, "Probabilistic diffusion of social influence with incentives," *IEEE Transactions on Services Computing*, vol. 7, no. 3, pp. 387–400, 2014.

[85] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[86] T. Kastrinogiannis, E.-E. Tsiropoulou, and S. Papavassiliou, "Utility-based uplink power control in cdma wireless networks with real-time services," in *International Conference on Ad-Hoc Networks and Wireless*, pp. 307–320, Springer, 2008.

[87] G. Katsinis, E. E. Tsiropoulou, and S. Papavassiliou, "Joint resource block and power allocation for interference management in device to device underlay cellular networks: A game theoretic approach," *Mobile Networks and Applications*, vol. 22, no. 3, pp. 539–551, 2017.

[88] H. Shakhatreh *et al.*, "Unmanned aerial vehicles (uavs): A survey on civil applications and key research challenges," *IEEE Access*, vol. 7, pp. 48572–48634, 2019.

[89] P. A. Apostolopoulos, M. Torres, and E. E. Tsiropoulou, "Satisfaction-aware data offloading in surveillance systems," in *Proceedings of the 14th Workshop on Challenged Networks*, pp. 21–26, 2019.

[90] M. Wan, G. Gu, W. Qian, K. Ren, X. Maldague, and Q. Chen, "Unmanned aerial vehicle video-based target tracking algorithm using sparse representation," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9689–9706, 2019.

[91] F. Vanegas, J. Roberts, and F. Gonzalez, "Uav tracking of mobile target in occluded, cluttered and gps-denied environments," in *2018 IEEE Aerospace Conference*, pp. 1–7, IEEE, 2018.

[92] Y. Liu, Q. Wang, H. Hu, and Y. He, "A novel real-time moving target tracking and path planning system for a quadrotor uav in unknown unstructured outdoor scenes," *IEEE Trans. on Syst., Man, and Cybern.: Syst.*, vol. 49, no. 11, pp. 2362–2372, 2018.

[93] W. Zhang, K. Song, X. Rong, and Y. Li, "Coarse-to-fine uav target tracking with deep reinforcement learning," *IEEE Trans. on Autom. Science and Eng.*, vol. 16, no. 4, pp. 1522–1530, 2018.

[94] A. Das, S. Shirazipourazad, D. Hay, and A. Sen, "Tracking of multiple targets using optimal number of uavs," *IEEE Trans. on Aerosp. and Electr. Syst.*, vol. 55, no. 4, pp. 1769–1784, 2018.

[95] N. H. Motlagh, M. Bagaa, and T. Taleb, "Uav-based iot platform: A crowd surveillance use case," *IEEE Communications Magazine*, vol. 55, no. 2, pp. 128–134, 2017.

[96] A. Farajzadeh, O. Ercetin, and H. Yanikomeroglu, "Uav data collection over noma backscatter networks: Uav altitude and trajectory optimization," in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pp. 1–7, IEEE, 2019.

[97] M. Samir, S. Sharafeddine, C. Assi, T. Nguyen, and A. Ghrayeb, "Uav trajectory planning for data collection from time-constrained iot devices," *IEEE Trans. on Wireless Communications*, vol. 19, no. 1, pp. 33–46, 2019.

[98] D. Gale and L. S. Shapley, "College admissions and the stability of marriage," *The American Mathematical Monthly*, vol. 69, no. 1, pp. 9–15, 1962.

[99] E. E. Tsiropoulou, P. Vamvakas, and S. Papavassiliou, "Joint utility-based uplink power and rate allocation in wireless networks: A non-cooperative game theoretic framework," *Physical Communication*, vol. 9, pp. 299–307, 2013.

[100] E. E. Tsiropoulou, G. K. Katsinis, and S. Papavassiliou, "Utility-based power control via convex pricing for the uplink in cdma wireless networks," in *2010 European Wireless Conference (EW)*, pp. 200–206, IEEE, 2010.

[101] K. Rael, G. Fragkos, J. Plusquellic, and E. E. Tsiropoulou, "Uav-enabled human internet of things," in *2020 16th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 312–319, IEEE, 2020.

[102] M. Mozaffari, W. Saad, M. Bennis, Y.-H. Nam, and M. Debbah, "A tutorial on uavs for wireless networks: Applications, challenges, and open problems," *IEEE communications surveys & tutorials*, vol. 21, no. 3, pp. 2334–2360, 2019.

[103] P.-V. Mekikis, A. Antonopoulos, E. Kartsakli, A. S. Lalos, L. Alonso, and C. Verikoukis, "Information exchange in randomly deployed dense wsns with wireless energy harvesting capabilities," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 3008–3018, 2016.

[104] M. Diamanti, E. E. Tsiropoulou, and S. Papavassiliou, "The joint power of noma and reconfigurable intelligent surfaces in swipt networks," in *2021 IEEE 22nd International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 621–625, IEEE, 2021.

[105] Z. Li, Y. Wang, M. Liu, R. Sun, Y. Chen, J. Yuan, and J. Li, "Energy efficient resource allocation for uav-assisted space-air-ground internet of remote things networks," *IEEE Access*, vol. 7, pp. 145348–145362, 2019.

[106] X. Lin, G. Su, B. Chen, H. Wang, and M. Dai, "Striking a balance between system throughput and energy efficiency for uav-iot systems," *IEEE IoT Journal*, vol. 6, no. 6, pp. 10519–10533, 2019.

[107] K. Zhu, X. Xu, and Z. Huang, "Energy-efficient routing algorithms for uav-assisted mmtc networks," in *IEEE 30th PIMRC*, pp. 1–6, IEEE, 2019.

[108] J. Xu, Y. Zeng, and R. Zhang, "Uav-enabled wireless power transfer: Trajectory design and energy region characterization," in *2017 IEEE Globecom Workshops*, pp. 1–7, IEEE, 2017.

[109] J. Xu, Y. Zeng, and R. Zhang, "Uav-enabled wireless power transfer: Trajectory design and energy optimization," *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 5092–5106, 2018.

[110] J. Park, H. Lee, S. Eom, and I. Lee, "Uav-aided wireless powered communication networks: Trajectory optimization and resource allocation for minimum throughput maximization," *IEEE Access*, vol. 7, pp. 134978–134991, 2019.

[111] P. Ramezani and A. Jamalipour, "Toward the evolution of wireless powered communication networks for the future internet of things," *IEEE network*, vol. 31, no. 6, pp. 62–69, 2017.

[112] P. Vamvakas, E. E. Tsiropoulou, M. Vomvas, and S. Papavassiliou, "Adaptive power management in wireless powered communication networks: A user-centric approach," in *2017 IEEE 38th Sarnoff Symposium*, pp. 1–6, IEEE, 2017.

[113] G. Baldini, S. Karanasios, D. Allen, and F. Vergari, "Survey of wireless communication technologies for public safety," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 2, pp. 619–641, 2013.

[114] D. Sikeridis, E. E. Tsiropoulou, M. Devetsikiotis, and S. Papavassiliou, "Wireless powered public safety iot: A uav-assisted adaptive-learning approach towards energy efficiency," *Journal of Network and Computer Applic.*, vol. 123, pp. 69–79, 2018.

[115] A. A. Khuwaja, Y. Chen, N. Zhao, M.-S. Alouini, and P. Dobbins, "A survey of channel modeling for uav communications," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 2804–2821, 2018.

[116] G. Fragkos, E. E. Tsiropoulou, and S. Papavassiliou, "Disaster management and information transmission decision-making in public safety systems," in *2019 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–6, IEEE, 2019.

[117] E. E. Tsiropoulou, T. Kastrinogiannis, and S. Papavassiliou, "Uplink power control in qos-aware multi-service cdma wireless networks," *Journal of Communications*, vol. 4, no. 9, pp. 654–668, 2009.

[118] Y. Wu, L. Qiu, and J. Xu, "Uav-enabled wireless power transfer with directional antenna: A two-user case," in *15th ISWCS*, pp. 1–6, IEEE, 2018.

[119] E. E. Tsiropoulou, A. Kapoukakis, and S. Papavassiliou, "Energy-efficient subcarrier allocation in sc-fdma wireless networks based on multilateral model of bargaining," in *2013 IFIP Networking Conference*, pp. 1–9, IEEE, 2013.

[120] K. Bhagat and J. Malhotra, "A survey of uplink multiple access techniques in lte mobile communication system," in *2014 International Conference on Advances in Engineering Technology Research (ICAETR - 2014)*, pp. 1–4, 2014.

[121] E. E. Tsiropoulou, A. Kapoukakis, and S. Papavassiliou, "Uplink resource allocation in sc-fdma wireless networks: A survey and taxonomy," *Computer Networks*, vol. 96, pp. 1–28, 2016.

[122] P. A. Thomas and M. Mathurakani, "Sc-fdma-an efficient technique for papr reduction in uplink communication systems-a survey," *International Journal of Research in Engineering and Technology*, vol. 3, no. 1, pp. 53–59, 2014.

[123] G. Fragkos, P. A. Apostolopoulos, and E. E. Tsiropoulou, "Escape: Evacuation strategy through clustering and autonomous operation in public safety systems," *Future Internet*, vol. 11, no. 1, p. 20, 2019.

[124] V. Chamola, V. Hassija, V. Gupta, and M. Guizani, "A comprehensive review of the covid-19 pandemic and the role of iot, drones, ai, blockchain, and 5g in managing its impact," *IEEE Access*, vol. 8, pp. 90225–90265, 2020.

[125] S. Swayamsiddha and C. Mohanty, "Application of cognitive internet of medical things for covid-19 pandemic," *Diabetes & Metabolic Syndrome: Clinical Research & Reviews*, 2020.

[126] Z. Schneeweiss, D. Murtaugh, and Bloomberg-Economics., "This is how deeply the coronavirus changed our behavior."

[127] M. M. Queiroz, D. Ivanov, A. Dolgui, and S. F. Wamba, "Impacts of epidemic outbreaks on supply chains: mapping a research agenda amid the covid-19 pandemic through a structured literature review," *Annals of Oper. Research*, pp. 1–38, 2020.

[128] S. Parthasarathy, S. Berryhill, C. Morton, A. Dean, A. Berryhill, N. Provencio-Dean, S. Patel, L. Estep, S. Mashaqi, D. Combs, *et al.*, "Effect of wearables on sleep in healthy individuals: A randomized cross-over trial and validation study," pp. A6165–A6165, American Thoracic Society, 2020.

[129] C. Burns., "Estimote wearables track workers to curb covid- 19 outbreak."

[130] Life-Signals, "Covid-19 remote health monitoring in hospitals and at home.."

[131] F. Khalique, S. A. Khan, and I. Nosheen, "A framework for public health monitoring, analytics and research," *IEEE Access*, vol. 7, pp. 101309–101326, 2019.

[132] R. Y. Kim, "The impact of covid-19 on consumers: Preparing for digital sales," *IEEE Engineering Management Review*, 2020.

[133] Mckinsey and Co., "Survey: Us consumer sentiment during the coronavirus crisis, mckinsey and co., april 5, 2020."

[134] W. J. McKibbin and R. Fernando, "The global macroeconomic impacts of covid-19: Seven scenarios," 2020.

[135] R. S. Sutton, A. G. Barto, *et al.*, *Introduction to reinforcement learning*, vol. 135. MIT press Cambridge, 1998.

[136] Y. Zhou, F. R. Yu, J. Chen, and Y. Kuo, "Cyber-physical-social systems: A state-of-the-art survey, challenges and opportunities," *IEEE Communications Surveys Tutorials*, vol. 22, no. 1, pp. 389–425, 2020.

[137] A. Thanou, E.-E. Tsiropoulou, and S. Papavassiliou, "A roadmap to congestion management in museums from a socio-technical perspective.," in *CI@ SMAP*, 2019.

[138] L. G. Jaimes, I. J. Vergara-Laurens, and A. Raij, "A survey of incentive techniques for mobile crowd sensing," *IEEE Internet of Things Journal*, vol. 2, no. 5, pp. 370–380, 2015.

[139] H. Gao, C. H. Liu, W. Wang, J. Zhao, Z. Song, X. Su, J. Crowcroft, and K. K. Leung, "A survey of incentive mechanisms for participatory sensing," *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 918–943, 2015.

[140] Z. Su, Q. Qi, Q. Xu, S. Guo, and X. Wang, "Incentive scheme for cyber physical social systems based on user behaviors," *IEEE Transactions on Emerging Topics in Computing*, vol. 8, no. 1, pp. 92–103, 2020.

[141] Y. Lin, Z. Chen, X. Feng, H. Zheng, and Y. Xu, "Incentive mechanism design for participatory sensing: Considering task quality and users' effort," in *2017 3rd IEEE International Conference on Computer and Communications (ICCC)*, pp. 707–711, 2017.

[142] Z. Akhtar, K. Siddique, A. Rattani, S. L. Lutfi, and T. H. Falk, "Why is multimedia quality of experience assessment a challenging problem?," *IEEE Access*, vol. 7, pp. 117897–117915, 2019.

[143] S. Cuomo, P. De Michele, A. Galletti, and F. Piccialli, "A cultural heritage case study of visitor experiences shared on a social network," in *2015 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC)*, pp. 539–544, 2015.

[144] R. P. Kinsley and J. Portenoy, "Perspectives of emerging museum professionals on the role of big data in museums," in *2015 48th Hawaii International Conference on System Sciences*, pp. 2075–2084, 2015.

[145] 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.

[146] M. G. Rashed, R. Suzuki, T. Yonezawa, A. Lam, Y. Kobayashi, and Y. Kuno, "Tracking visitors in a real museum for behavioral analysis," in *2016 Joint 8th International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems (ISIS)*, pp. 80–85, 2016.

[147] M. G. Rashed, D. Das, Y. Kobayashi, and Y. Kuno, "Analysis and prediction of real museum visitors' interests and preferences based on their behaviors," in *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, pp. 451–456, 2017.

[148] 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.

[149] 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.

[150] I. Lykourentzou, X. Claude, Y. Naudet, E. Tobias, A. Antoniou, G. Lepouras, and C. Vassilakis, "Improving museum visitors' quality of experience through intelligent recommendations: A visiting style-based approach.," in *Intelligent environments (workshops)*, pp. 507–518, 2013.

[151] 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.

[152] E. Véron and M. Levasseur, *Ethnographie de l'exposition: l'espace, le corps et le sens*. Bibliothèque publique d'information du Centre Pompidou, 1989.

[153] G. Hafeez, K. S. Alimgeer, Z. Wadud, I. Khan, M. Usman, A. B. Qazi, and F. A. Khan, "An innovative optimization strategy for efficient energy management with day-ahead demand response signal and energy consumption forecasting in smart grid using artificial neural network," *IEEE Access*, vol. 8, pp. 84415–84433, 2020.

[154] A. Al Hadi, C. A. S. Silva, E. Hossain, and R. Challoo, "Algorithm for demand response to maximize the penetration of renewable energy," *IEEE Access*, vol. 8, pp. 55279–55288, 2020.

[155] R. Zhou, Z. Li, and C. Wu, "An online procurement auction for power demand response in storage-assisted smart grids," in *IEEE INFOCOM*, pp. 2641–2649, IEEE, 2015.

[156] A. Kumari, R. Gupta, S. Tanwar, S. Tyagi, and N. Kumar, "When blockchain meets smart grid: Secure energy trading in demand response management," *IEEE Network*, vol. 34, no. 5, pp. 299–305, 2020.

[157] D. Li, W.-Y. Chiu, H. Sun, and H. V. Poor, "Multiobjective optimization for demand side management program in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1482–1490, 2017.

[158] H. Wang and J. Huang, "Incentivizing energy trading for interconnected microgrids," *IEEE Transactions on Smart Grid*, vol. 9, no. 4, pp. 2647–2657, 2018.

[159] A. Nikoobakht, J. Aghaei, M. Shafie-Khah, and J. P. S. Catal̃áo, "Assessing increased flexibility of energy storage and demand response to accommodate a high penetration of renewable energy sources," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 2, pp. 659–669, 2019.

[160] T. Morstyn and M. D. McCulloch, "Multiclass energy management for peer-to-peer energy trading driven by prosumer preferences," *IEEE Transactions on Power Systems*, vol. 34, no. 5, pp. 4005–4014, 2019.

[161] S. L. Arun and M. P. Selvan, "Intelligent residential energy management system for dynamic demand response in smart buildings," *IEEE Systems Journal*, vol. 12, no. 2, pp. 1329–1340, 2018.

[162] E. S. Kang, S. J. Pee, J. G. Song, and J. W. Jang, "A blockchain-based energy trading platform for smart homes in a microgrid," in *2018 3rd Int. Conf. on Computer and Communication Systems*, pp. 472–476, 2018.

[163] W. Liu, J. Zhan, and C. Y. Chung, "A novel transactive energy control mechanism for collaborative networked microgrids," *IEEE Transactions on Power Systems*, vol. 34, no. 3, pp. 2048–2060, 2019.

[164] H.-H. Chang, W.-Y. Chiu, H. Sun, and C.-M. Chen, "User-centric multiobjective approach to privacy preservation and energy cost minimization in smart home," *IEEE Systems Journal*, vol. 13, no. 1, pp. 1030–1041, 2019.

[165] Y. Liu, L. Xiao, G. Yao, and S. Bu, "Pricing-based demand response for a smart home with various types of household appliances considering customer satisfaction," *IEEE Access*, vol. 7, pp. 86463–86472, 2019.

[166] A. Safdarian, M. Fotuhi-Firuzabad, and M. Lehtonen, "Optimal residential load management in smart grids: A decentralized framework," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1836–1845, 2016.

[167] C. Yang and W. Lou, "On optimizing demand response management performance for microgrids under communication unreliability constraint," in *IEEE GLOBECOM*, pp. 1–6, IEEE, 2015.

[168] C. Yang, J. Yao, W. Lou, and S. Xie, "On demand response management performance optimization for microgrids under imperfect communication constraints," *IEEE Internet of Things Journal*, vol. 4, no. 4, pp. 881–893, 2017.

[169] S. Wang, S. Bi, and Y.-J. A. Zhang, "The impacts of energy customers demand response on real-time electricity market participants," in *IEEE Int. Conf. on Communications (ICC)*, pp. 1–7, IEEE, 2018.

[170] S. Wang, S. Bi, and Y.-J. A. Zhang, "Demand response management for profit maximizing energy loads in real-time electricity market," *IEEE Transactions on Power Systems*, vol. 33, no. 6, pp. 6387–6396, 2018.

[171] L. Jia and L. Tong, "Dynamic pricing and distributed energy management for demand response," *IEEE Transactions on Smart Grid*, vol. 7, no. 2, pp. 1128–1136, 2016.

[172] USA EIA, "U.S. Energy Information Administration," 2021.

Appendix A

Author Publications

1. **N. Patrizi**, E.E. Tsiropoulou, and S. Papavassiliou, "Health Data Acquisition from Wearable Devices during a Pandemic: A Techno-Economics Approach," in IEEE ICC, 2021. (to appear)

2. **N. Patrizi**, G. Fragkos, E.E. Tsiropoulou, and S. Papavassiliou, "Contract - Theoretic Resource Control in Wireless Powered Communication Public Safety Systems," in IEEE GLOBECOM, pp. 1-6, 2020.

3. **N. Patrizi**, G. Fragkos, K. Ortiz, M. Oishi, and E.E. Tsiropoulou, "A UAV-enabled Dynamic Multi-Target Tracking and Sensing Framework," in IEEE GLOBECOM, pp. 1-6, 2020.

4. G. Fragkos, **N. Patrizi**, E. E. Tsiropoulou, and S. Papavassiliou, "Socio-aware Public Safety Framework Design: A Contract Theory based Approach," ICC 2020 - 2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, pp. 1-7, 2020.

5. **N. Patrizi**, P.A. Apostolopoulos, K. Rael, and E.E. Tsiropoulou, "Socio-physical Human Orchestration in Smart Cities," in IEEE International Conference on Smart Computing (SMARTCOMP), pp. 115-120, 2019.

6. **N. Patrizi**, S.K. LaTouf, E.E. Tsiropoulou, and S. Papavassiliou, "Museum and Visitor Interaction & Feedback Orchestration Enabled by Labor Economics," in IEEE Transactions on Computational Social Systems. (major revision)

7. **N. Patrizi**, S.K. LaTouf, E.E. Tsiropoulou, and S. Papavassiliou, "Prosumer-centric Self-sustained Smart Grid Systems" in IEEE Systems Journal (to appear)

Appendix B

Honors and Awards

Received the SMART (Science, Mathematics, and Research for Transformation) scholarship from the United States Department of Defence.