Fall 10-8-2020

Socio-physical Human Orchestration in Smart Cities

Nathan Patrizi

Nathan Patrizi

Candidate

Electrical and Computer Engineering

Department

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

Approved by the Thesis Committee:

Eirini Eleni Tsiropoulou, Chairperson

Payman Zarkesh-Ha

Jim Plusquelic

Socio-physical Human Orchestration in Smart Cities

by

Nathan Patrizi

B.S, University of New Mexico, 2019

THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science
Computer Engineering

The University of New Mexico

Albuquerque, New Mexico

December, 2020

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 and pushing me to succeed.

I would also like to thank Pavlos Athanasios Apostolopoulos and Kelly Rael for their help in performing the research presented in this thesis.

Socio-physical Human Orchestration in Smart Cities

by

Nathan Patrizi

B.S, University of New Mexico, 2019

M.S., Computer Engineering, University of New Mexico, 2020

Abstract

In this thesis, we present a novel socio-physical human orchestration framework to deal with the increasing complexity of smart city environments, by capitalizing on recent advances in game theory and reinforcement learning, with the goal of solving the efficient management of a smart city while improving the quality of life for the humans living in said smart city. This problem has become more challenging as the smart cities have become more complicated and complex with new technologies. In the proposed framework, each human selects a Point of Interest (PoI) that it wants to visit. This selection is performed by acting as stochastic learning automaton, which evaluates 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, through adopting the theory of minority games. The resulting Nash equilibrium point classifies which of the humans that finally visit each PoI. A low complexity reinforcement learning based algorithm is used achieve the overall framework. Finally, a detailed set of numerical

and comparative results are shown to provide information on the efficiency of the approach.

This work has been published in:

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

Contents

List of Figures

Glossary

N	Set of humans where N = $\{1, ..., n, ..., N\}$
S	Set of Points of Interest where S = $\{1, ..., s, ..., S\}$
$i_{n,s}$	Personal interest of human n in visiting PoI s
$SI_{n,j}$	Willingness of human n and j to socially interact
c_s	cost of PoI s serving a human
N_s^{thres}	Physical capacity of the PoI s
$\|N\|_s^{Go}$	Number of humans who go to PoI s
$d_{n,s}$	distance of human n from PoI s
$QoS_{n,s}$	Quality of Service experienced by human n at PoI s
$r_{n,s}$	Reward of the human n to visist PoI s
$Pr_{n,s}$	Probability human n visits PoI s
A_n	Set of strategies for a human n to go or not go, denoted as $A_n = 0, 1$
a_n	Human n specific strategy $a_n \in A_n$
f_{a_n}	Payoff function for human n based on their action a_n

Chapter 1

Overview

1.1 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 [1]. 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.

1.2 Background & Motivation

1.2.1 Game Theory & Reinforcement Learning in Cyber-Physical Soscial Systems

Smart cities act as cyber-physical social systems consisting of components related to control, communications, and computing [2, 3]. Several recent research works in the field of cyber-physical social systems exploit and study the Quality of Experience of the end-users in dynamically changing environments, such as art places [4, 5, 6], mobile communications systems [7, 8], electronic communication services and applications [9], museums [10, 11, 12, 13], public safety events in smart cities [14, 15, 16, 17], man-made disaster management scenarios [18, 19, 20, 21], and others. The aforementioned research works adopt distributed solutions and methods, such as game theory [22] and reinforcement learning [23, 24], in order to address the corresponding problem of optimizing the end-users Quality of Experience given the nature of the cyber-physical social systems, where lack of information is observed among the involved entities in the system.

Game theory is a powerful tool in order to capture the preferences and the corresponding decisions of end-users with opposite and competitive interests within a dynamically changing environment [25]. Game Theory has been applied in several fields and applications of the cyber-physical social systems, focusing on the control, communications, and computing aspects of them, such as the unmanned aerial vehicles-based communications [26, 27] and computing applications [28, 29] to enable the efficient coordination of the different types of entities, sensors, humans, etc., residing within the smart city's environment.

Reinforcement learning is a category of machine learning focusing on how the autonomous agents take actions in an environment in order to maximize the notion of a cumulative reward. Reinforcement learning is one of three basic machine learning paradigms, alongside supervised learning and unsupervised learning [30]. Focusing on the field of cyber-physical social systems, reinforcement learning techniques have been applied in smart grid systems to enable the autonomous decision making of the end-users regarding their electricity consumption, the choice of the utility company to be served from [31, 32], as well as the pricing policy making [33], in dynamic wireless environments in order to enable the end-users to select the provider to be served from [34, 35, 36], in smart museums to enable the time and cost efficient visitors' touring [37, 38], just to name some indicative applications.

Based on the above discussion, it is envisioned that the tools of game theory and reinforcement learning are becoming part of the artificial intelligence era in order to enable the smart and distributed decision making within complex environments [39, 40]. Indeed, Artificial Intelligence (AI) has emerged as a powerful tool to support devices' autonomous human-like decision-making, while being founded on and supported by multi-disciplinary techniques, such as machine learning, control theory, game theory, optimization theory, and meta-heuristics [41, 42].

1.2.2 Smart City Planning & Points of Interests Visits

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 [43], Meetup [44], and Twitter [45, 46, 47]. In [48], 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 [49], 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

Figure 1.1: Smart City Illustration and Sensitive Services (Source: Internet of Business)

groups' behavioral patterns. In [50], 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 [51], 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 [52], 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

[53], 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 [54] 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 [55] 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 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.

1.3 Smart City Networks Applications, Challenges, and Future Problems

1.3.1 Applications in smart cities

The first application considered is the EBSNs previously mentioned. These newer social networks have similar functionality to that of traditional social networks, but with the added aspect of allowing the user to communicate their location and attendance to an event to other users. This broadcasting of user's location and status allows for some new opportunities to facilitate a better experience when using an EBSN. Most notably, businesses and public spaces (parks, national monuments, museums, etc.) can use this data to learn more about where people like to spend time and get a better idea for what events individuals are actively participating in towards improving the designed recommender systems and improving the end users' perceived Quality of Experience [56, 57, 58]. This allows for those businesses to adapt to what people like or offer them an incentive to get them to visit their business [59, 60].

Another application space is in crowd sourcing real time geography based information. These applications are based around the users providing their own specific information to the service provider, which then uses the information to improve their own services. A prominent example is the traffic data that apple and google maps provide. This information is provided to google and apple from the users, which is then used to construct the traffic information that is presented to the users in the application. The users are compensated for this information by being able to use the feature.

1.3.2 Challenges in Smart cities

Clearly, these new systems will have a number of associated challenges. These would range from a pure performance standpoint to facilitating individuals choosing to adopt these new system into their everyday life. We first analyze some of the potential problems that these systems face from a pure networking point of view. The networking of these systems

Some users will have an inherent resistance to changing how they interact to include these newer solutions into their lives. This could be due to concerns about privacy or just being unwilling to change. With the work going into developing and providing these options to the public, it is important that they choose to adopt them. This leads to the problems of incentive providing. One solution of which was presented above in Section 1.3.1 with the example of crowd-sourcing navigation data. This approach works well as users have an inherent desire to participate in that system as it could reduces travel times and avoid potentially congested areas. This approach would work well in such cases were the service provided inherently relies on the information provided from each user to create an accurate model, as a more accurate model will provide a better result for all the users to benefit from. However, this wouldn't necessarily apply if the user doesn't naturally gain from the improvement of the model. Consider a case where a business would like to know more information about the needs of the humans in the community. These humans may be resistant to providing this information due to any number of concerns. To overcome this, the business should incentive the humans to provide the information such that the humans become willing to do so.

1.3.3 Future problems in smart cities

As networks become more and more complex, it will continue to be important to improve upon the fundamental infrastructure that these communication systems operate on. With the internet of things' impact on smart cities as a whole, it will naturally share many of the same fundamental problems. Aspects such as interference management and efficient power consumption will become increasingly important in this field as the number of devices will continue to increase which will need to communicate effectively.

Efficient management of the smart city is another fundamental problem. Clearly, as more information becomes available from all of the various systems in smart cities, smart cities should use that information to coordinate various services and provide a better experience for all citizens that live in the smart city.

The last open problem is the need to have some dynamicity in a smart city network. A truly smart city would be able to know about potential future events that could impact the performance of the system and adapt to perform better with this knowledge. This type of adaptation is becoming more important as further improvements in network performance continue.

1.4 Contributions

This thesis 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.

1.5 Outline

The rest of the thesis is organized as follows. In Section 2.1, the overall system model is described. While in Section 2.2 the proposed human-centric reinforcement learning-based PoI selection process is presented in detail. Section 2.3 introduces the autonomous human orchestration to the PoIs based on the theory of minority games, additionally in Section 2.4 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 Chapter 3. Specifically, the pure performance is evaluated in Section 3.2 and comparative results are presented in Section 3.3. Finally, Chapter 4 provides the conclusion to this thesis.

Chapter 2

Human Orchestration in Smart cities

2.1 System Model

In this thesis, 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 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

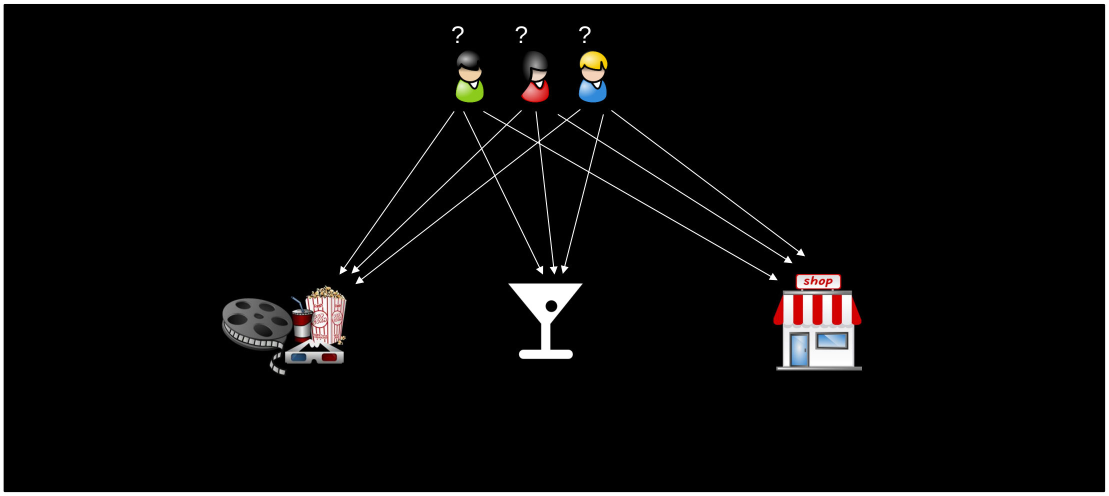

Figure 2.1: Humans autonomous decision making regrading the Point of Interest that they will visit.

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 $|N|_s^{Go} \leq N_s^{thres}$.

2.2 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 [31] 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 2.1): (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

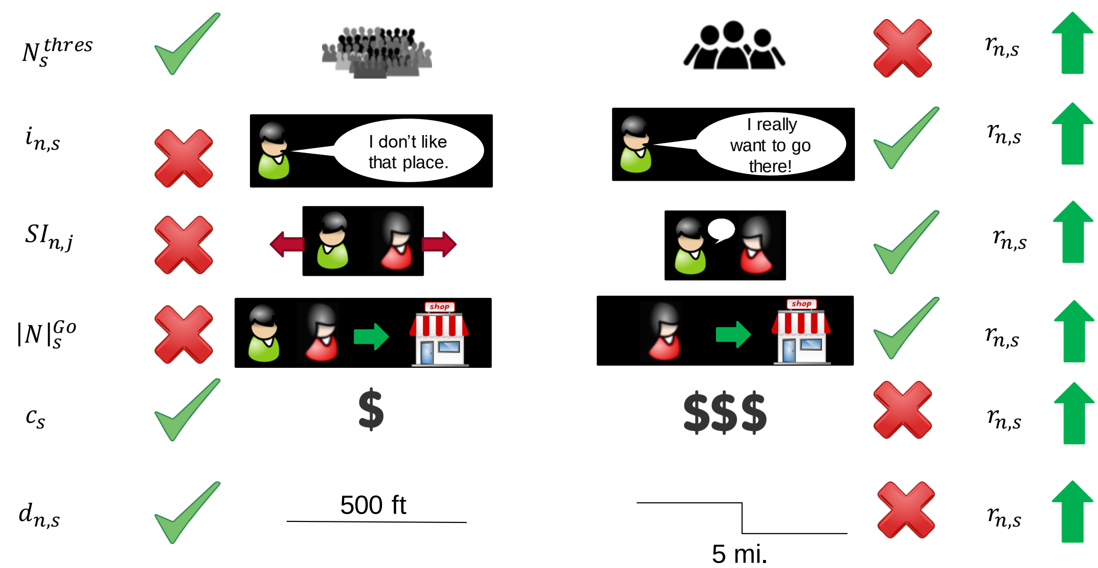

Figure 2.2: Graphical representation of the reward function.

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

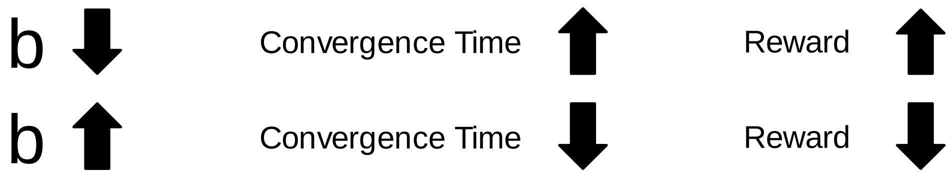

Figure 2.3: Graphical representation of the dependence of the convergence time and the corresponding achieved reward from the learning rate b.

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 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 [31, 61, 62], the humans update their action probabilities based on the following rule [63].

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

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

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 2.4.

2.3 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) [31].

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 2.2). 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 thesis, 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 2.4).

2.4 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 $\mathbf{Pr_n^{[t]}} = [Pr_{n,1}^{[t]}, \ldots, Pr_{n,|S|}^{[t]}]$. 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 [64]. 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 $f^{ite}_{a_n}$ (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 $\tilde{r}^{[t+1]}_{n,s}, \forall s \in S$, and updates its action probabilities $Pr^{[t+1]}_{n,s}, \forall s \in S$ (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, 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.

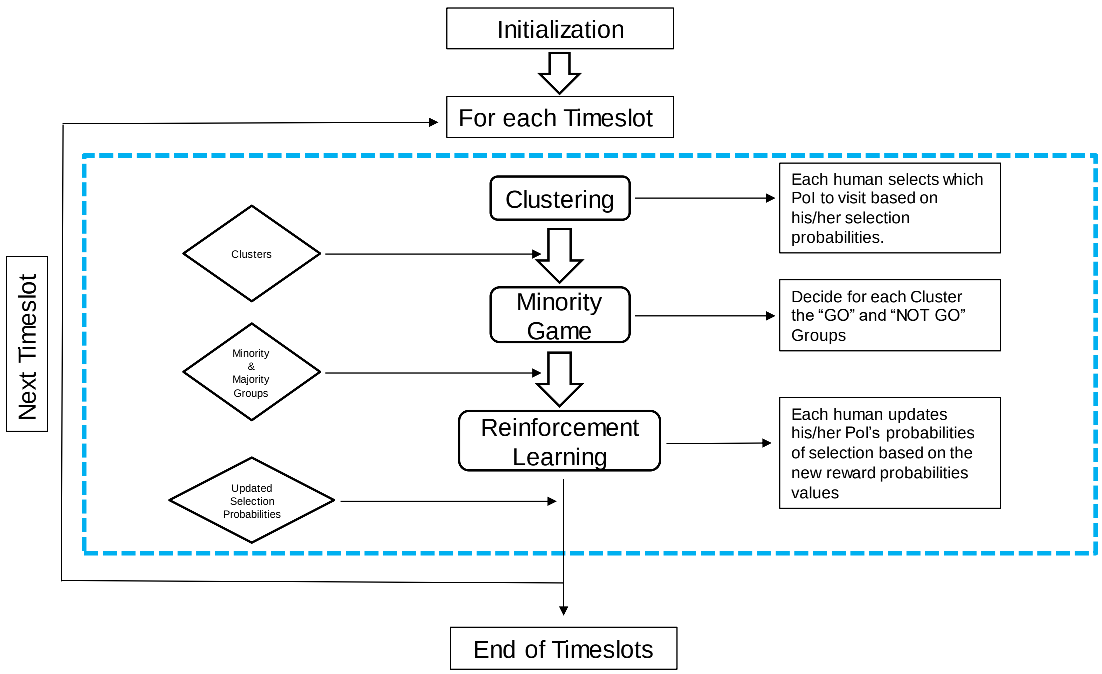

Figure 2.4: Graphical representation of the overall proposed framework in this research work summarizing the flow of information, as well as of the control actions to conclude to the autonomous decision making process.

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: $\mathbf{Pr}^* = [\mathbf{Pr}_\mathbf{1}^*, \ldots, \mathbf{Pr}_\mathbf{n}^*, \ldots, \mathbf{Pr}_{|\mathbf{N}|}^*]$
while $Conv == 0$ **do**
  $N_s = \varnothing, \forall s \in S$
  Choose $a_n(t) \in S, N_{a_n(t)} = N_{a_n(t)} \cup \{n\}, \forall n \in N$
  based on $\mathbf{Pr}_\mathbf{n}^{[\mathbf{t}]}$
  $MinorityGame(N_s, N_s^{thres}), \forall s \in S$
  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* $\forall n \in N, \forall s \in S$
  $Conv = 1,$ ***if*** $\forall n \in N, \exists s \in S$: $|Pr_{n,s}^{[t+1]} - 1| \leq 0.99$
  $t = t + 1$
end while
$\mathbf{Pr}_\mathbf{n}^* = \mathbf{Pr}_\mathbf{n}^{[\mathbf{t}]}, \forall n \in N$
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**
  $N_s^{GO} = N_s^{NGO} = \varnothing$
  Choose $a_n^{ite}, \forall n \in N_s$ *based on* $pr_n^{ite} = [pr_{n,0}^{ite}, pr_{n,1}^{ite}]$
  if $a_n^{ite} = 1,$ ***then*** $N_s^{GO} = N_s^{GO} \cup \{n\}$
  else $N_s^{NGO} = N_s^{NGO} \cup \{n\}$
  if $|N_s^{GO}| \leq N_s^{thres},$ ***then*** $w^{ite} = 1$ ***else*** $w^{ite} = 0$
  $\pi_{n,a_n}^{ite+1} = \pi_{n,a_n}^{ite} + f_{a_n}^{ite}$
  $pr_{n,a_n}^{ite+1} = \exp(\gamma \cdot \pi_{n,a_n}^{ite+1}) / \sum_{\forall a_n \in A_n} \exp(\gamma \cdot \pi_{n,a_n}^{ite+1})\ \forall a_n \in A_n, \forall n \in N_s$
  $Conv = 1,$ ***if*** $\forall n \in N_s, \exists a_n \in A_n$: $|pr_{n,a_n}^{ite+1} - 1| \leq 0.99$
  $ite = ite + 1$
end while
EndFunction

Chapter 3

Experiments

3.1 Experiment Setup

In this chapter, a detailed numerical evaluation of the proposed approach is presented in terms of the overall framework's operation efficiency (Section 3.2) and superiority compared to other alternatives (Section 3.3). 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.

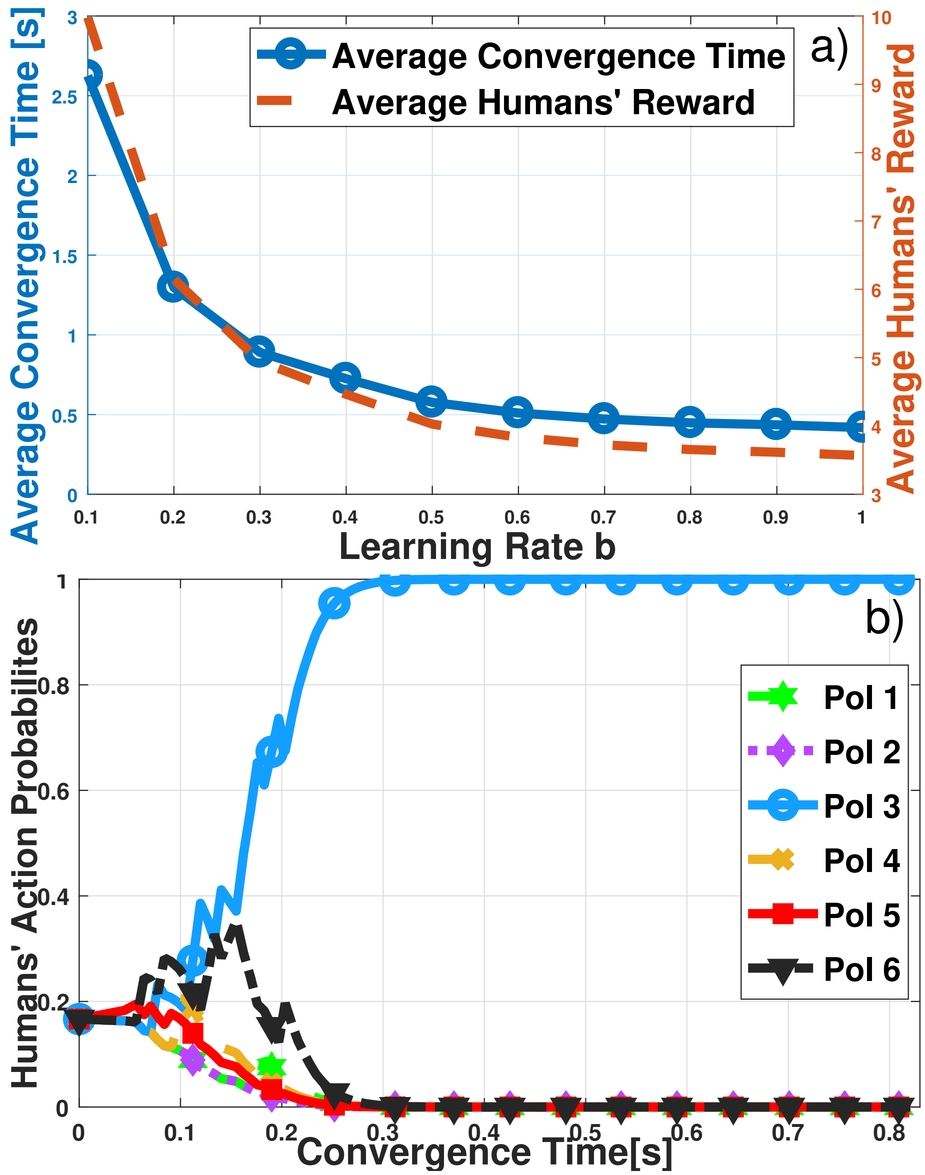

Figure 3.1: (a) Average convergence time and average humans' reward vs b and (b) Action probabilities convergence

3.2 SmartPoI Framework's Operation

First, we evaluate the operation of the socio-physical PoI selection following the proposed reinforcement learning technique. Figure 3.1a 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 3.1b, 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 3.2a-3.2d, 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 3.2 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 3.2(a)-3.2(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 3.2a) and the higher capacity (Figure 3.2c). The results also illustrate that the humans select the PoI with the closest physical proximity (Figure 3.2b). In Figure 3.2d 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.

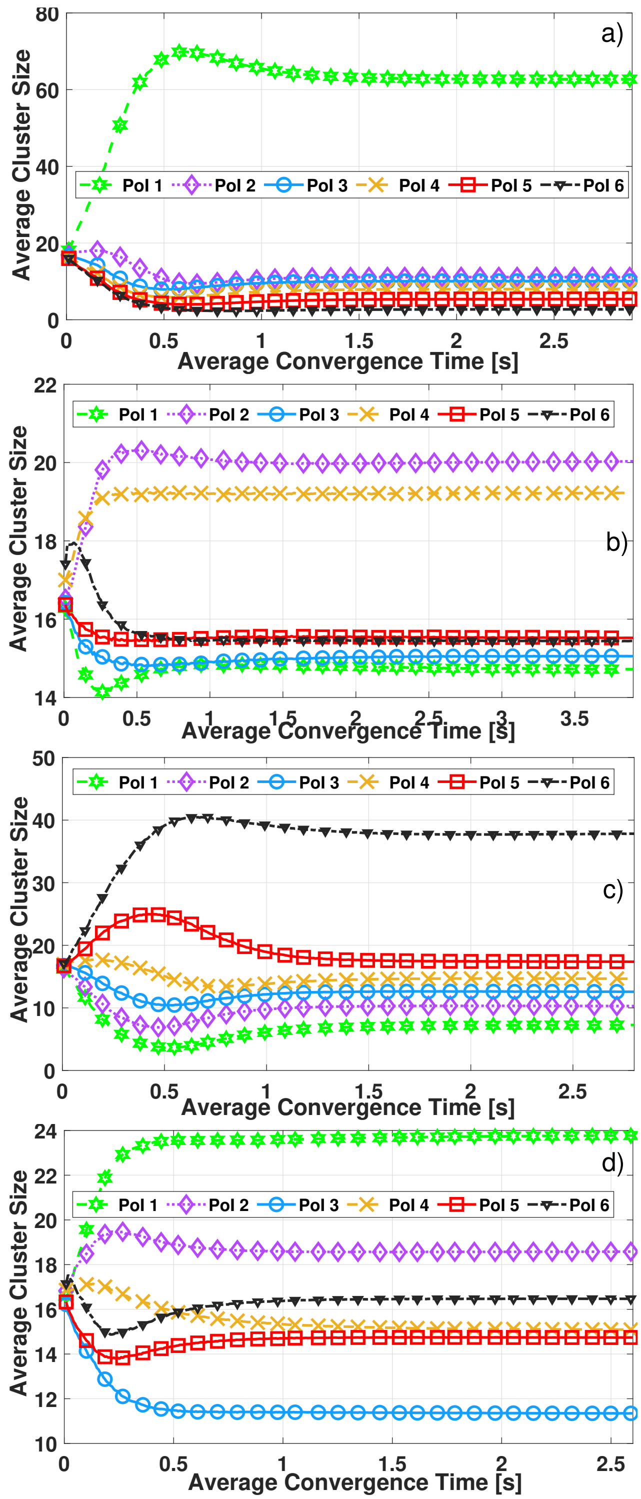

Figure 3.2: Humans' cluster size per PoI for varying (a) PoIs' cost c_s (b) humans' distance from PoIs' $d_{n,s}$, (c) PoIs' capacity N_s^{thres}, and (d) all the socio-physical factors in Eq.2.1

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 3.3a for two indicative subjects. Also, Figure 3.3b 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.

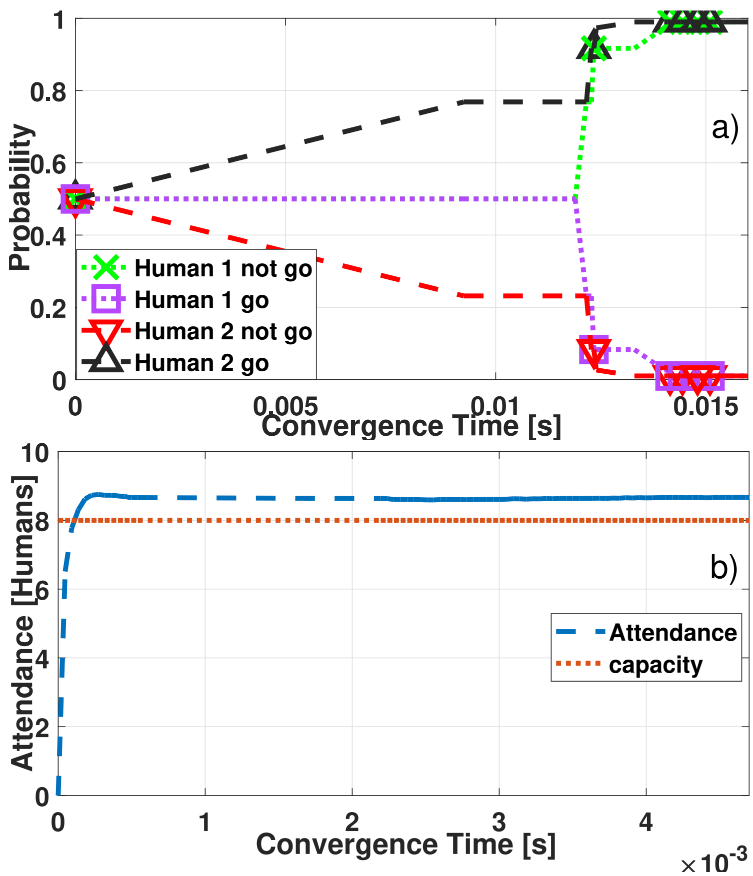

Figure 3.3: Convergence of human's (a) action probabilities and (b) attendance.

3.3 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 3.4a). Also the average convergence time to the PoIs selection (Figure 3.4b) and the average cluster size of humans per PoI (Figure 3.4c) 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 3.4a), while allowing the humans to quickly learn their desired PoI selection (Figure 3.4b) and not overcongest the PoIs (Figure 3.4c).

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.

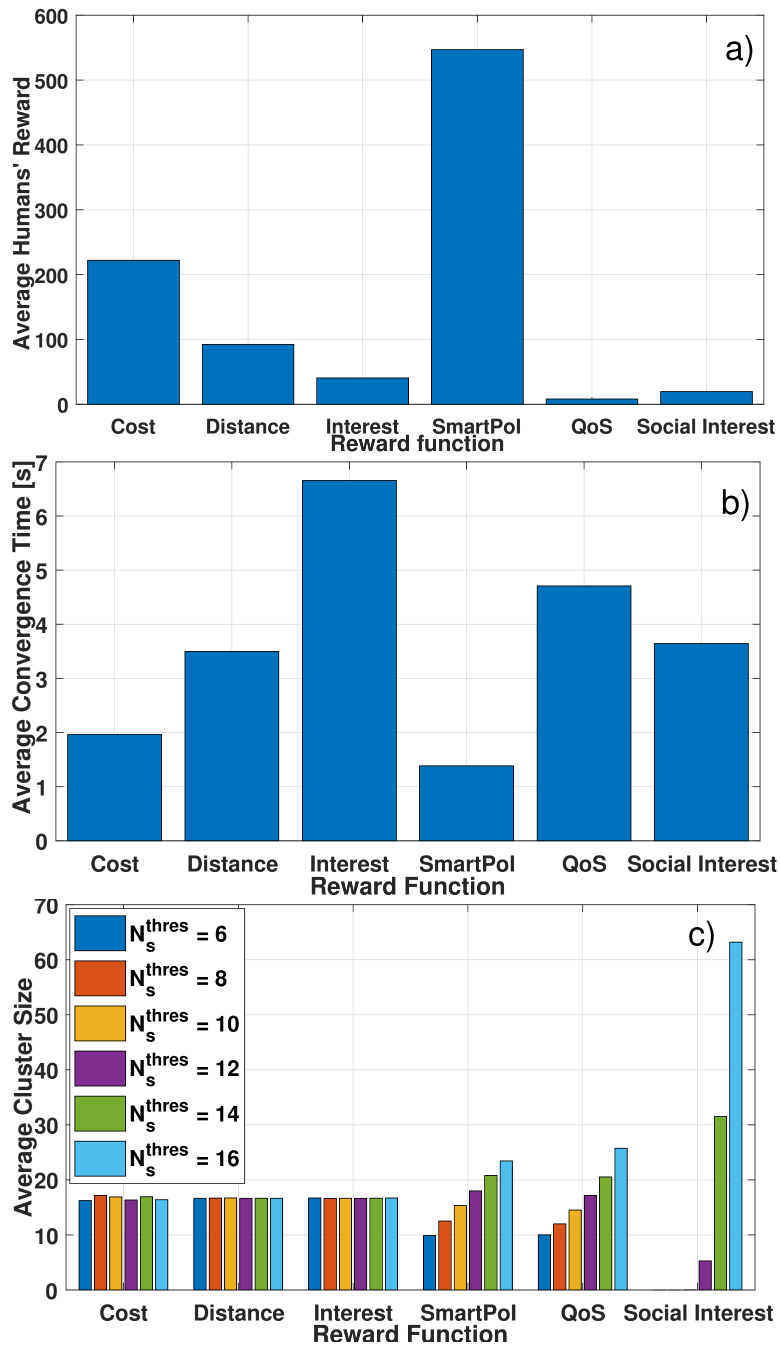

Figure 3.4: Average (a) Humans reward (b) Convergence time, and (c) Cluster size per PoI, for different reward functions

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 3.4c 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 3.5a), 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}.

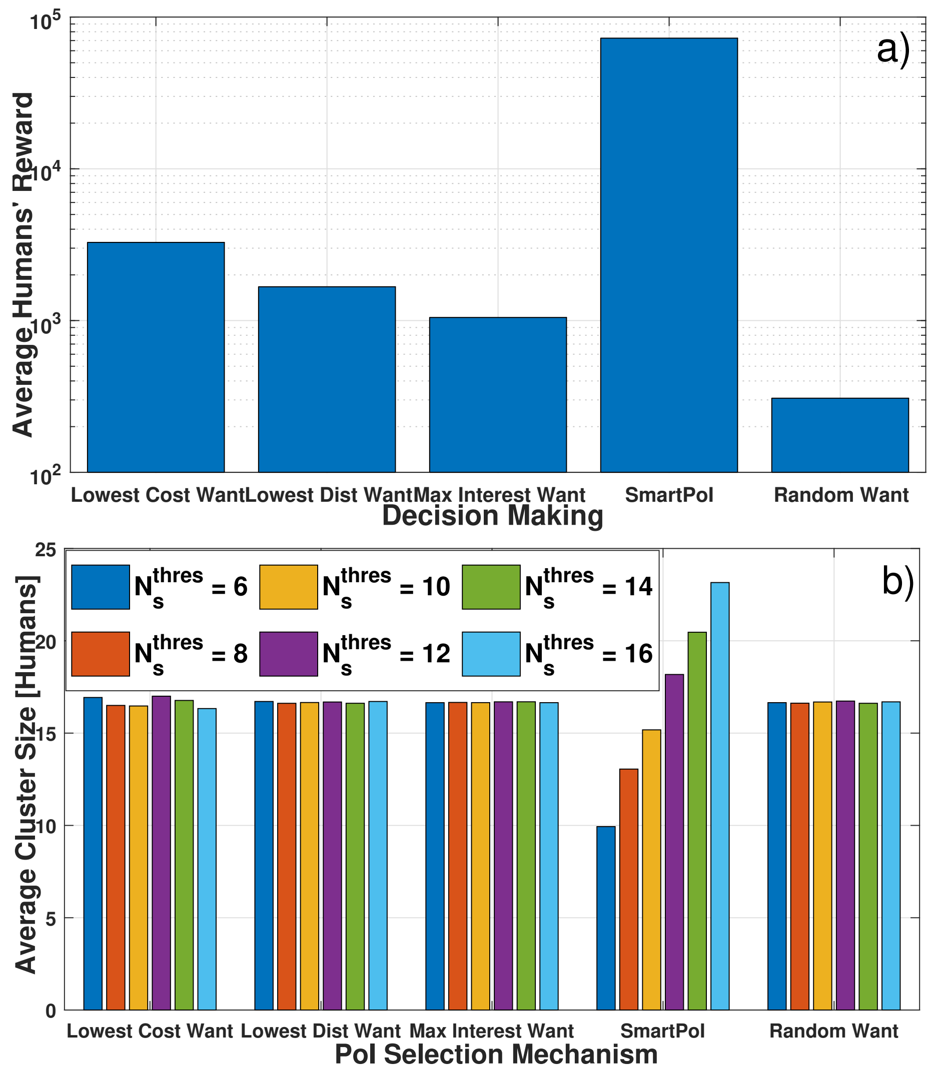

Figure 3.5: Average (a) Humans reward and (b) cluster size per PoI, for different PoIs selection mechanism

Chapter 4

Conclusion and Future Works

In this thesis, 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.

References

[1] 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.

[2] C. G. Cassandras, "Smart cities as cyber-physical social systems," *Engineering*, vol. 2, no. 2, pp. 156–158, 2016.

[3] 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.

[4] 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.

[5] 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.

[6] E. Shmueli, V. K. Singh, B. Lepri, and A. Pentland, "Sensing, understanding, and shaping social behavior," *IEEE Transactions on Computational Social Systems*, vol. 1, no. 1, pp. 22–34, 2014.

[7] M. Dong, T. Kimata, K. Sugiura, and K. Zettsu, "Quality-of-experience (qoe) in emerging mobile social networks," *IEICE TRANSACTIONS on Information and Systems*, vol. 97, no. 10, pp. 2606–2612, 2014.

[8] K. Kilkki, "Quality of experience in communications ecosystem.," *J. UCS*, vol. 14, no. 5, pp. 615–624, 2008.

[9] U. Reiter, K. Brunnström, K. De Moor, M.-C. Larabi, M. Pereira, A. Pinheiro, J. You, and A. Zgank, "Factors influencing quality of experience," in *Quality of experience*, pp. 55–72, Springer, 2014.

[10] 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.

[11] P. Brooks and B. Hestnes, "User measures of quality of experience: why being objective and quantitative is important," *IEEE network*, vol. 24, no. 2, pp. 8–13, 2010.

[12] 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.

[13] 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.

[14] 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.

[15] D. He, S. Chan, and M. Guizani, "Drone-assisted public safety networks: The security aspect," *IEEE Communications Magazine*, vol. 55, no. 8, pp. 218–223, 2017.

[16] R. Favraud, A. Apostolaras, N. Nikaein, and T. Korakis, "Toward moving public safety networks," *IEEE Communications Magazine*, vol. 54, no. 3, pp. 14–20, 2016.

[17] 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.

[18] 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.

[19] R. Verdone and S. Mignardi, "Joint aerial-terrestrial resource management in uav-aided mobile radio networks," *IEEE Network*, vol. 32, no. 5, pp. 70–75, 2018.

[20] V. Gunes, S. Peter, T. Givargis, and F. Vahid, "A survey on concepts, applications, and challenges in cyber-physical systems.," *KSII Transactions on Internet & Information Systems*, vol. 8, no. 12, 2014.

[21] P. Vamvakas, E. E. Tsiropoulou, and S. Papavassiliou, "On the prospect of uav-assisted communications paradigm in public safety networks," in *IEEE INFOCOM 2019-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)*, pp. 762–767, IEEE, 2019.

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

[23] X.-L. Huang, X. Ma, and F. Hu, "Machine learning and intelligent communications," *Mobile Networks and Applications*, vol. 23, no. 1, pp. 68–70, 2018.

[24] L. P. Kaelbling, M. L. Littman, and A. W. Moore, "Reinforcement learning: A survey," *Journal of artificial intelligence research*, vol. 4, pp. 237–285, 1996.

[25] D. Fudenberg and J. Tirole, "Game theory, 1991," *Cambridge, Massachusetts*, vol. 393, no. 12, p. 80, 1991.

[26] P. Vamvakas, E. E. Tsiropoulou, and S. Papavassiliou, "Exploiting prospect theory and risk-awareness to protect uav-assisted network operation," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, no. 1, p. 286, 2019.

[27] M. E. Mkiramweni, C. Yang, J. Li, and W. Zhang, "A survey of game theory in unmanned aerial vehicles communications," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 4, pp. 3386–3416, 2019.

[28] 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.

[29] 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.

[30] E. Alpaydin, *Introduction to machine learning*. MIT press, 2020.

[31] 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.

[32] R. Lu, S. H. Hong, and X. Zhang, "A dynamic pricing demand response algorithm for smart grid: reinforcement learning approach," *Applied Energy*, vol. 220, pp. 220–230, 2018.

[33] R. Lu and S. H. Hong, "Incentive-based demand response for smart grid with reinforcement learning and deep neural network," *Applied energy*, vol. 236, pp. 937–949, 2019.

[34] P. Vamvakas, E. E. Tsiropoulou, and S. Papavassiliou, "Dynamic provider selection & power resource management in competitive wireless communication markets," *Mobile Networks and Applications*, vol. 23, no. 1, pp. 86–99, 2018.

[35] O. Naparstek and K. Cohen, "Deep multi-user reinforcement learning for dynamic spectrum access in multichannel wireless networks," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pp. 1–7, IEEE, 2017.

[36] C. Zhang, Z. Liu, B. Gu, K. Yamori, and Y. Tanaka, "A deep reinforcement learning based approach for cost-and energy-aware multi-flow mobile data offloading," *IEICE Transactions on Communications*, p. 2017CQP0014, 2018.

[37] T. Misu, K. Georgila, A. Leuski, and D. Traum, "Reinforcement learning of question-answering dialogue policies for virtual museum guides," in *Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, pp. 84–93, 2012.

[38] 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.

[39] Y. Lu and K. Yan, "Algorithms in multi-agent systems: A holistic perspective from reinforcement learning and game theory," *arXiv preprint arXiv:2001.06487*, 2020.

[40] Z. Gao, "Understanding the future of deep reinforcement learning from the perspective of game theory," in *Journal of Physics: Conference Series*, vol. 1453, p. 012076, 2020.

[41] R. Li, Z. Zhao, X. Zhou, G. Ding, Y. Chen, Z. Wang, and H. Zhang, "Intelligent 5g: When cellular networks meet artificial intelligence," *IEEE Wireless communications*, vol. 24, no. 5, pp. 175–183, 2017.

[42] G. Fragkos, E. E. Tsiropoulou, and S. Papavassiliou, "Artificial intelligence enabled distributed edge computing for internet of things applications," in *2020 16th International Conference on Distributed Computing in Sensor Systems (DCOSS)*, pp. 450–457, IEEE Computer Society, 2020.

[43] T. H. Silva, P. O. Vaz de Melo, J. M. Almeida, J. Salles, and A. A. Loureiro, “A comparison of foursquare and instagram to the study of city dynamics and urban social behavior,” in *Proceedings of the 2nd ACM SIGKDD international workshop on urban computing*, pp. 1–8, 2013.

[44] H. Ding, C. Yu, G. Li, and Y. Liu, “Event participation recommendation in event-based social networks,” in *International conference on social informatics*, pp. 361–375, Springer, 2016.

[45] X. Liu, Q. He, Y. Tian, W.-C. Lee, J. McPherson, and J. Han, “Event-based social networks: linking the online and offline social worlds,” in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1032–1040, 2012.

[46] A. Q. Macedo, L. B. Marinho, and R. L. Santos, “Context-aware event recommendation in event-based social networks,” in *Proceedings of the 9th ACM Conference on Recommender Systems*, pp. 123–130, 2015.

[47] P. Giridhar, S. Wang, T. Abdelzaher, T. Al Amin, and L. Kaplan, “Social fusion: Integrating twitter and instagram for event monitoring,” in *2017 IEEE International Conference on Autonomic Computing (ICAC)*, pp. 1–10, IEEE, 2017.

[48] 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.

[49] 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.

[50] 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.

[51] 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.

[52] 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.

[53] 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.

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

[55] 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.

[56] 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.

[57] A. Q. De Macedo and L. B. Marinho, "Event recommendation in event-based social networks.," in *HT (Doctoral Consortium/Late-breaking Results/Workshops)*, Citeseer, 2014.

[58] 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.

[59] H. Yin, L. Zou, Q. V. H. Nguyen, Z. Huang, and X. Zhou, "Joint event-partner recommendation in event-based social networks," in *2018 IEEE 34th International Conference on Data Engineering (ICDE)*, pp. 929–940, IEEE, 2018.

[60] Y. Liu, A. Liu, X. Liu, and X. Huang, "A statistical approach to participant selection in location-based social networks for offline event marketing," *Information Sciences*, vol. 480, pp. 90–108, 2019.

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

[62] 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.

[63] 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.

[64] 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.