5-2018

Role of opinion sharing on the emergency evacuation dynamics

Aravinda Ramakrishnan Srinivasan
University of Tennessee, asriniv2@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Aravinda Ramakrishnan Srinivasan entitled "Role of opinion sharing on the emergency evacuation dynamics." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Mechanical Engineering.

Subhadeep Chakraborty, Major Professor

We have read this dissertation and recommend its acceptance:

William R. Hamel, Asad J. Khattak, Jindong Tan, Eric R. Wade

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Role of opinion sharing on the emergency evacuation dynamics

A Dissertation Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Aravinda Ramakrishnan Srinivasan

May 2018

To my mom,

Revathi Srinivasan

my sisters,

Nithya and Krithika

my teachers,

and my friends

Acknowledgements

First and foremost, I would like to thank Dr. Subhadeep Chakraborty for believing in me and being a major support during my graduate studies in the University of Tennessee, Knoxville. The freedom he gave for us to think and propose new ideas, his insightful comments and always encouraging attitude has gone a long way in helping me cope with the pressure of graduate school.

Next I would like to thank my colleagues, Farshad Salimi Naneh Karan, Christoper LaBorde, Russell Graves, David Harris, Adam Whitfield, Navneet Kikkeri for giving awesome support during all the ups and downs one can face during the graduate school. I would have been long lost without them teaching me the American culture, innumerable witty comments and silly pep talks.

Further I would like to name some of my friends who would take my phone call no matter what time of the day it is to listen to my troubles, give some encouraging words and most importantly be there for me when I was feeling low! Thank you Raghuraman, Ajay Krishnan, Ashwin Govindan and Jagannathan for bearing with me.

I would also like to appreciate my family, Revathi (Mom), Nithya (Sister) and Krithika (Sister) for being there for the past 25+ years. They lived through my constant tantrums and never ceased to love and support me.

Last but not the least, I would like to thank my committee members, Dr. William Hamel, Dr. Jingdong Tan, Dr. Eric Wade and Dr. Asad Khattak for being supportive during my graduate school. Words can not express my gratitude to my committee.

"We are what our thoughts have made us;

so take care about what we think"

- Swami Vivekananda

Abstract

Emergency evacuation is a critical research topic and any improvement to the existing evacuation models will help in improving the safety of the evacuees. Currently, there are evacuation models that have either an accurate movement model or a sophisticated decision model. Individuals in a crowd tend to share and propagate their opinion. This opinion sharing part is either implicitly modeled or entirely overlooked in most of the existing models. Thus, one of the overarching goal of this research is to the study the effect of opinion evolution through an evacuating crowd. First, the opinion evolution in a crowd was modeled mathematically. Next, the results from the analytical model were validated with a simulation model having a simple motion model. To improve the fidelity of the evacuation model, a more realistic movement and decision model were incorporated and the effect of opinion sharing on the evacuation dynamics was studied extensively. Further, individuals with strong inclination towards particular route were introduced and their effect on overall efficiency was studied. Current evacuation guidance algorithms focuses on efficient crowd evacuation. The method of guidance delivery is generally overlooked. This important gap in guidance delivery is addressed next. Additionally, a virtual reality based immersive experiment is designed to study factors affecting individuals' decision making during emergency evacuation.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction and Literature Review

1.1 Introduction

Emergency evacuation is a stressful situation and the safety of all occupants is of prime concern to building planners. Emergency can be triggered due to several reasons: natural hazards such as fire, earthquake, etc. or man-made emergencies, such as active shooters, stampede, etc. Since 1982, there have been at least 81 public mass shootings across the USA, with the killings occurring in 33 states from Massachusetts to Hawaii [1]. Forty-four of these mass shootings have occurred since 2006. Seven of them took place in 2012 alone, including Sandy Hook. An analysis of this database by researchers at Harvard University, further corroborated by a FBI study, determined that mass shootings have been on the rise. Similarly, there were 1.346 million fires in the U.S. with 3280 deaths and $14.3 billion loss in 2015 alone [2]. In response to this alarming trend, emergency evacuation of buildings has been identified as an important topic of research. Optimization of pedestrian flow can possibly decrease the time spent along non-optimal paths and hence reduce damage related to panic situations. However, such optimization processes are challenging since crowds need to be interpreted not only as an assembly, but also as individuals who aggregate or disaggregate according to specific strategies.

The planning authorities take into account various factors when deciding on a particular evacuation procedure for a given building in case of an emergency. Factors like maximum capacity of an exit, maximum allowed occupancy of the building, minimum time to evacuate,

etc. play important role in this planning. Researchers have tried to model the evacuation procedure to study the various parameters involved and to optimize the evacuation plan. Several existing emergency evacuation simulators try to take into account as many factors as possible to effectively calculate the time to evacuate and also test the efficacy of different evacuation procedures or to compute an optimized plan for evacuation.

A comprehensive literature review of the state-of-the-art in pedestrian and emergency evacuation research is provided here to motivate the research carried out in this dissertation and also to serve as a quick resource for researchers in the field.

1.2 Literature Review

1.2.1 Fluid flow models

Hughes et al. [3, 4, 5] modeled the movement of pedestrians as fluid flow by identifying the governing equations for the fluid flow model. This work identifies the human crowd as analogous to thinking fluids and studies the effect of barrier placement to improve the flow of the crowd. Colombo et al. [6] studied the effect of panic with the continuum fluid flow model. These models, though conforming largely with experimental data fails to take into account the role of interactions among egressing individuals in determining their exit choices and overall movement dynamics.

1.2.2 Cellular automata models

Dijkstra et al. [7] and Blue et al. [8] presented cellular automata (CA) models for pedestrian flow. Each node was given a set of rules and the emergent behavior of the society was found to be valid with existing macro-level observations. Burstedde et al. [9] and Kirchner et al. [10] introduced a stochastic cellular automaton model with a static and a dynamic floor field combined. The dynamic floor field was analogous to chemotaxis, every node would leave a trail that diffused and decayed at a specific rate. The static floor field accounted for the attraction towards exit, repulsion from obstacles and other constant forces. The dynamic field accounted implicitly for communication between the nodes. Kirchner et al. [11] studied

the effect of model parameters on the overall behavior of the system. Nishinari et al. [12] presented the ant trail model and the floor field model and showed the similarity between them.

Krichner et al. [13] investigated a cellular automata model with respect to competitive and cooperative behavior by including a friction coefficient while combining the floor fields. Kluepfel [14] presented a complete study of how different velocities of individuals can be computed taking into account various factor like age, gender, etc. and modeled the competition between people as analogous to Newtonian friction. Henein et al. [15] added the concept of interpersonal force, the force exerted by individuals on one another in a crowded environment to account for injuries and studied the effect of it on the evacuation time. Varas et al. [16] delved into effect of obstacles, the effect of door size and position of the door on the egress model in a cellular automata (CA) world. Shiwakoti et al. [17] used ant society to learn model parameters and employed a scaling concept from biology to scale the parameters for a human society. They studied the effect of structural features/layout on the egress dynamics. Seitz et al. [18] discussed the problem of losing line of sight to the leader in a leader-follower scenario in CA environment. Alizadeh [19] presented an improved CA model with a dynamic floor field that considered the density of the crowd at exits.

Though these models have their unique way of representing the motion of individuals, there is scope to improve the decision mechanism that supports the motion model. In these works, the exit choice were generally determined by the static floor field and crowd impatience with their current exit choice were not incorporated. However, they incorporated a collision avoidance and an implicit opinion sharing through the dynamic floor field.

1.2.3 Social force based models

Helbing et al. [20] introduced the concept of the social force model. In the social force model, an individual's movement is affected by factors like their desired velocity, tendency to maintain minimum distance from others, and attractive force of an exit. The effects of the environment and the crowd were captured intrinsically in this model.

Parisi et al. [21] used the social force model to study the effect of different degrees of panic. The different degrees of panic is simulated through different desired individual velocity. The

effect of door sizes with different panic level on the evacuation time is studied in detail. The panic level is found to affect the formation of clusters among the nodes and the distribution of cluster mass/size is found to have to a 'U' shaped characteristic curve with the panic level/desired velocity [22]. Again, this model gives a comprehensive simulation for only the movement dynamics of evacuation.

Zhou et al. [23] modeled crowd evacuation in the presence of an aggressive attacker with a fuzzy inference system. Their model was similar to the social force model and there were no direct interpersonal interactions. Helbing et al. presented a detailed summary of the existing body of work on emergent systems with focus on behavioral models in [24, 25].

1.2.4 Lattice models

Lattice gas models have been used to model and verify a classroom evacuation in [26]. The particles in the simulation execute a biased random walk toward the exit and the model takes into account personal space/minimum distance to avoid collision as well as obstacles, but in this work, the exit choice was predetermined and lacked an explicit model for route choices. Takimoto et al. [27] used the lattice gas model of pedestrian movement to study the relationship between escape time and the starting position of people from a room with single exit. Additionally, the effect of exit width on the distribution of escape time was examined. Song et.al and Guo et al. [28, 29] combined the lattice gas model with the social force model. They tried to incorporate interaction among individuals implicitly through the force fields. The average evacuation time found using simulations combining both models were found to be more accurate compared to the lattice gas model alone. This further strengthens the need for a decision making model that more explicitly takes into account exit choice as well as one-on-one and group interactions.

1.2.5 Discrete event models and game theoretic models

Lino et al. [30] modeled the crowd egress dynamics with the principles of queuing in networks. Singh et al. [31] utilized a discrete event model. The effect of leaders and sub groups in crowd dynamics was examined in detail, but did not involve inter-personal opinion sharing

to support their movement model. Lo et al. [32] used a game theoretic approach to model the exit choice of individuals. A virtual agent played against each individual till Nash equilibrium was arrived for each step to decide on the exit choice for that individual.

1.2.6 Discrete choice models

Bierlaire et al. [33] presented a technical report on utilizing discrete choice model to predict and keep track of pedestrians for automatic video surveillance. Antonini et al. [34] introduced a discrete choice model for predicting a pedestrian's instantaneous decision. The destination and route were known and the model parameters were calibrated from real walking data to predict the next position for a given pedestrian. They performed utility maximization to predict the pedestrian's choice and their potential application was tracking pedestrian in video surveillance. The model was improved by adding kinematic leader-follower and collision-avoidance pattern and validated with recorded data in [35]. Hoogendoorn et al. [36] delved into pedestrian route choice modeling when they have several activities to perform. The effect of other individuals were implicitly taken into account by accounting for density of people in the desired route. A utility maximization was used to find the desired route of a pedestrian. It is evident that the effect of opinion sharing is largely absent in these techniques.

Lovreglio et al. [37, 38] utilized stated preference of surveyed individuals to fit a random utility model for the herding behavior of evacuees during evacuation. This approach has the same shortcoming as previous models since individuals sharing their opinions during emergency evacuation is absent from their model.

1.2.7 Other simulation models

Pan et al. [39, 40, 41] presented a multi-agent simulation engine. It incorporated a decision tree architecture to model different behavior of evacuating individuals. Behaviors like leader following, group member seeking their group were inculcated into their model by Chu et al. in [42] highlighting the importance of incorporating social behavior. Pelechano et al. [43] presented a graph search based simulation with trained/untrained leaders (know the entire

map/not) and a crowd of followers and found out that having a small percentage of trained leader can lead to quicker evacuation. It does not account for individuals' impatience and the movement model is relative simple. Korhonen et al. [44] presented an agent-based modeling technique with different types of agents. The agents did not have impatience incorporated into them but they showed that agents actively searching for exits were able to help the agents that were passively following.

1.2.8 Psychological models

Proulx [45] stresses the need for better understanding of human interaction under emergency. Hasan et al. [46, 47] examined the effect of a person's social network on their decision to evacuate after receiving a hurricane warning. It was found that individuals' social links and the amount of trust they have on their links strongly influences their decision to evacuate. Goldstone et al. [48] used agent based modeling to study the group behavior from a psychology point of view but this lacks the complementary motion model to become a complete evacuation model. Spieser et al. [49, 50, 51, 52] studied just the psychological dynamics in opinion control. Using Gustav LeBon's suggestibility theory [53], a discrete-time non linear model of crowd psychological behavior was developed. The elements of a queue were agitated and a control algorithm to bring the agitated elements to normal state using one or more control node(s) was derived. Though the psychological aspect is well modeled, lack of a motion model limits its utility as a complete evacuation model.

1.2.9 Experimental models

Drury et al. [54] presented a virtual reality (VR) based computer simulation of evacuating an underground train station. This paper incorporated VR to investigate the psychological aspect of emergency evacuation. Nicolas et al. [55] studied experimentally the effect of selfish behavior in a crowd of polite people on the flow rate through an exit.

Moussaïd et al. [56] collected data from a set of well-controlled experiments to understand the laws governing pedestrian behavior during simple avoidance tasks. Moussaïd et al. [57] presented a simple heuristic vision-based behavior model. Pedestrians were modeled to

minimize deviation from a straight line path to exit while avoiding obstacles. Collision with other pedestrians and walls were accounted for in the force equation and they found that this simple behavior-based model matched well with experimental data. In another work, Moussaïd et al. [58] attempted to study the effect of social interaction on egress dynamics within a virtual environment. The social interactions were limited to line-of-sight and the herding effect and bottlenecks were studied in detail.

Bode and Codling [59] investigated crowd behavior in a 2D virtual environment. The results have to be taken with a grain of salt since the 2D environment provided a bird's eye view of the entire environment. Bode et al. [60] extended their study and investigated the effect of different information sources on individual's exit choice by utilizing the same virtual setup. They concluded that different information sources combined had a unique effect when compared to individual sources.

1.2.10 Guidance models

Work has been done to model and simulate a crowd guidance mechanism to help the crowd to safety in shortest possible time. Gao et al. [61] take into account the confidence level of individuals in either accepting or rejecting guidance. Wang et al. [62] tried to build an intelligent crowd guidance system by giving a probability of accepting the guidance. Directed graphs and Markov Decision Processes were used to solve the optimization with respect to avoiding blocked pathway. This work also incorporated model for fire propagation and crowd impatience in the optimization. These works highlight the importance for a guidance system to help optimize the evacuation time.

Kuligowski [63] and Zheng et al. [64] gave comprehensive overview of existing simulation techniques. Also, Duives et al. [65] presented a comprehensive review of existing literature on crowd movement and classified them according to the model capabilities.

Averill [66] identified some challenges in the field of pedestrian and evacuation dynamics. He pointed out the need for a stochastic model of the evacuation process and adoption of technology for collecting data and modeling human behavior. Also, he stressed out the need for theoretical behavior model with numerical validity.

From the extensive review of literature the following gaps are found and the research objectives are designed to alleviate the shortcomings of the existing techniques.

1.3 Gaps

Gap 1: There is a lack of analytical model to study the effect of opinion propagation through an evacuating crowd. The two complementary dynamics of decision sharing and movement has not been analytically and numerically investigated in the context of emergency evacuation. Additionally, it is also possible that a few amongst the crowd will display a strong attraction towards one of the exit choices available, borne out of prior knowledge of the environment or confidence in their decision making. They can propagate their opinion extensively due to their natural predisposition to be leaders. The effect of strong opinion holders on the evacuation dynamics poses an interesting problem.

Gap 2: A careful study of the existing literature indicates a lack of parametric study on the herding behavior and various factors affecting the decision making process. If an individual becomes impatient quickly with bottlenecks and changes their route, then does it affect the overall egress dynamics? If an individual ponders over his/her decision frequently and adjust their decision, then will it help the crowd to evacuate faster? Does having leader(s) in the crowd helpful or detrimental to the evacuation process? Does having individuals who are more receptive to others' opinion improves the efficiency?

Gap 3: Current state-of-the-art guidance systems do not explicitly take into account how the guidance is delivered to individuals/groups. Due to their predisposition, some individuals might try to follow the guidance, some might just imitate neighbors disregarding the guidance provided, while some like security personnel, building manager,etc. might actively propagate the guidance issued by the central guidance system. It is not practically to assume that everyone will follow the optimal evacuation route without active instructions. A guidance algorithm that takes into account this variability in crowd nature and provides an active guidance delivery instruction to a responsible individual is lacking in current literature.

1.4 Research objectives

Objective 1 *(addressing gap 1)*: **Mathematical modeling and numerical validation of decision sharing.** In the first phase of the project, the combined movement and decision sharing model will be defined using stochastic differential equation. The time evolution of the opinion (decision) propagation through the crowd will be mathematically derived from the stochastic equation. A corresponding simulation model will be developed and the effect of leaders or strong opinionated individuals on the overall distribution of the crowd at the available exits will be investigated.

Objective 2 *(addressing gap 2)* : **A Markov decision process based decision theoretic model with spatially bounded opinion sharing framework.** In the second stage of the project, a Markov decision process based decision model will be defined. The model will take into account the collision avoidance behavior of individuals, impatience exhibited by crowd at bottlenecks, re-evaluation of current route choice by evacuees at regular intervals, and herding behavior through a spatially bounded opinion sharing framework. A thorough parametric study of the factors affecting the overall efficiency of the evacuation process will be conducted.

Objective 3 *(addressing gap 3)* : **A dynamic guidance algorithm.** The lessons learned through implementation of objective 2 will be accounted for in a dynamic guidance algorithm. The guidance system will monitor current status of evacuation and update the guidance provided to the evacuees to improve the overall efficiency of the process. This will result in a realistic estimate of evacuation time for a given building structure. Also, a guide can be given a set of active instructions to navigate through the building to optimize the evacuation time of the evacuees.

Objective 4 *(addressing gap 3)* : **A virtual reality set up to validate the factors affecting individuals' exit choice.** Having developed a guidance model and a set of instruction for a guide to improve the emergency evacuation procedure, the final task will

involve developing an immersive virtual reality set up to test and validate the factors affecting individuals' decision. Virtual reality platform serves as an excellent tool to test out emergency evacuation scenarios since it does not put any individual through a real-life threatening situation. Nevertheless, virtual reality platform will help to elicit realistic information when compared to written/oral survey techniques about individual's choices during emergency evacuation. This can lead to designing better delivery of guidance to the evacuating crowd, thus improving the overall evacuation of the crowd.

Chapter 2

Pedestrian dynamics with explicit sharing of exit choice during egress through a long corridor

2.1 Objective

A careful look at the state-of-the-art in egress literature shows that although tremendous progress has been made in modeling pedestrian movement in emergency, the effect of 'herding' tendencies on egress dynamics has not received as much attention. This part starts with a simple egress situation but incorporates the effect of group interaction on route choice and hence the movement dynamics of individuals *. The movement dynamics in turn affects the instantaneous formation and dissipation of small groups of evacuees. To the best of our knowledge, these two complementary dynamics (decision and movement) has not been analytically and numerically investigated in the context of egress in the past.

In an emergency, a group will make their choices of different escape routes by taking into account not only their individual predispositions, distances to exits, familiarity with the environment, obstacles in their path, perceived sense of danger, etc., but also through

*Results presented in this section are from our published work: Srinivasan, A. R., Karan, F. S. N., & Chakraborty, S. (2017). Pedestrian dynamics with explicit sharing of exit choice during egress through a long corridor. Physica A: Statistical Mechanics and its Applications, 468, 770-782.

imitation of and influence from people who are physically nearby. It is also possible that a few amongst the crowd will display a strong attraction towards one of the exit choices available, borne out of prior knowledge of the environment or confidence in their decision making. A few might spread their opinion extensively due to their natural predisposition to be leaders. The effect of strong opinion holders on the egress dynamics poses an interesting problem.

This work studies a simplified model of movement along with an opinion sharing framework to study the combined effect of both.

2.2 Modeling of crowd movement dynamics with opinion sharing

For the purpose of this work, a long corridor with two exits (an exit to left (E_L) and an exit to right (E_R)) is considered. At each instant, each individual of the crowd can choose to use either of the two exits and correspondingly, move one step toward right or left end exit of the corridor. To account for the explicit swapping of exit choice information among individuals, the voter model dynamics is utilized. According to this dynamics, at arbitrary time steps, one random individual is spontaneously influenced by one of his physically close neighbors, chosen at random. If the neighbor happens to be moving in the same direction as him, he finds reinforcement in his belief that he is indeed going in the optimal (safest) direction. If the neighbor happens to be rushing towards the other exit, that introduces doubt and leads to him changing his decision. Plausibly, this influence is modeled to get weaker as the individual and his neighbor are further away from each other. In the analytical model, all individuals are assumed to start at the center of the corridor and an interaction zone starting at the center of the corridor and stretching to 20% of the total length of the corridor on either side is established. The individuals can successfully affect other's exit decision only if both are within the interaction zone. To account for the motion model, after every decision step, every individual of the crowd moves one step towards their respective exit choice. However, if an individual changes his/her exit choice they are assumed to be

able to join up with the new group instantaneously. The rationale behind this assumption is borrowed from considering the crowd movement as a choked fluid flow through a narrow bottleneck, where people move much slower as a group due to crowding in a narrow space. Consequently, the passage between the two groups is largely empty and the individuals switching between groups can join their new group relatively quickly. To study this motion with opinion model analytically, a Master equation approach, developed previously in [67, 68] is utilized. For completeness and clarity, the master equation and a polynomial solution is derived below.

2.2.1 Analytical Solution

Let, at a given instant, the number of people without strong opinions moving toward E_R be denoted by N_R and the number moving toward E_L be denoted by N_L. The total number of indecisive people moving is thus $N = N_R + N_L$. In addition, there are I_R people strongly predisposed to move toward the right exit while I_L having a strong bias toward the left exit, for a total of $I = I_R + I_L$ evacuees with strong opinions. This is illustrated in Fig. 2.1

Let us now define three variables - the crowd polarization parameter, $p = \frac{N_R - N_L}{N}$, the influencer ratio, $u = \frac{I_R - I_L}{I}$ and the global influence ratio, $\zeta = \frac{I}{N}$. The crowd polarization parameter $p \in [-1, 1]$ captures the ratio of people moving right vs. moving left. Thus, $p = 1$ means that everybody is moving towards the right exit at that instant, $p = -1$ means that everybody is moving towards the left exit at that instant and $p = 0$ means that half are moving towards right and the rest toward left. The influencer ratio, u denotes the relative influence or control that people with strong opinions, (who we will subsequently identify as 'leaders') have over the independent decision makers' possible exit choices. $u = \pm 1$ denotes each of the independent thinkers are moving towards the exit on the right side (or left side)

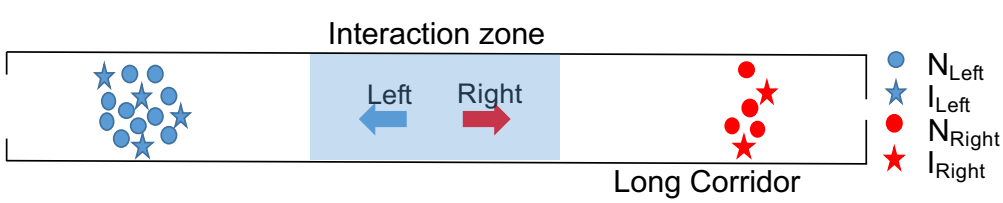


Figure 2.1: Illustration of a long narrow corridor with a group moving toward either side

of the long corridor, $u = 0$ indicates that there is equal number of leaders attracting the crowd towards both the exits. The global influence ratio, ζ is the fraction of the number of influencers to the number of indecisive people. $\zeta = 0$ implies there are no influencers in the crowd. As ζ increases, the number of strong opinion holders in the crowd increases until at $\zeta = 1$ the whole crowd comprises individuals holding strong opinions. The master equation for this stochastic system is given by

$$\dot{P}_p = r_{p+\frac{2}{N}} P_{p+\frac{2}{N}} + g_{p-\frac{2}{N}} P_{p-\frac{2}{N}} - (r_p + g_p) P_p \tag{2.1}$$

where,

$$\begin{aligned} r_p = P(p \to p - \frac{2}{N}) &= \left(\frac{N_R}{N}\right)\left(\frac{N_L + I_L}{N + I - 1}\right) \\ g_p = P(p \to p + \frac{2}{N}) &= \left(\frac{N_L}{N}\right)\left(\frac{N_R + I_R}{N + I - 1}\right) \\ r_{p+\frac{2}{N}} &= \left(\frac{N_R + 1}{N}\right)\left(\frac{N_L - 1 + I_L}{N + I - 1}\right) \\ g_{p-\frac{2}{N}} &= \left(\frac{N_L + 1}{N}\right)\left(\frac{N_R - 1 + I_R}{N + I - 1}\right) \end{aligned} \tag{2.2}$$

Substituting Eqn. 2.2 in Eqn. 2.1, we get

$$\begin{aligned} \dot{P}_p = &\left(\frac{N_R + 1}{N}\right)\left(\frac{N_L - 1 + I_L}{N + I - 1}\right) P_{p+\frac{2}{N}} \\ &+ \left(\frac{N_L + 1}{N}\right)\left(\frac{N_R - 1 + I_R}{N + I - 1}\right) P_{p-\frac{2}{N}} \\ &- \left[\left(\frac{N_R}{N}\right)\left(\frac{N_L + I_L}{N + I - 1}\right) + \left(\frac{N_L}{N}\right)\left(\frac{N_L + I_L}{N + I - 1}\right)\right] P_p \end{aligned} \tag{2.3}$$

For large N, assuming that $I < N$, $I_L P_{p+2/N} + I_R P_{p-2/N} \approx I P_p$, with proper scaling of time as $\tau = t/N^2$ and noting that $N_R I_L - N_L I_R = \frac{NI}{2}(p - i)$, the master equation can be simplified [67] to its final form as,

$$\frac{\partial P_p}{\partial \tau} = \frac{1}{2}\frac{\partial^2}{\partial p^2}\left[B(p) P_p\right] - \frac{\partial}{\partial p}\left[A(p) P_p\right] \tag{2.4}$$

$$\text{where,} \quad B(p) = 2\left(1 - p^2\right) \tag{2.5}$$

$$A(p) = I\left(u - p\right) \tag{2.6}$$

Equation 2.4 describing the time evolution of the probability density function of the polarization parameter p, can be recognized as the Fokker-Planck equation and can be treated with generic methods developed for such partial differential equations. Wong et.al. [69] has reported certain general conditions under which the problem reduces to an eigenvalue problem of the Sturm-Liouville type and gives rise to polynomial solutions. If it is assumed that an equilibrium density function exists, and

$$\lim_{\tau \to \infty} \frac{\partial P_p}{\partial \tau} = 0 \tag{2.7}$$

then it is simple to show that the equilibrium density $p_e(m)$ satisfies

$$\frac{d}{dp}\left((1 - p^2)p_e(p)\right) - I(u - p)p_e(p) = 0 \tag{2.8}$$

if the constants of integration are assumed to be 0. Substituting $P_p(\tau) = f(\tau)p_e(p)\varphi(p)$, in Eqn. 2.4 and using separation of variables,

$$\frac{df(\tau)}{d\tau} = -\lambda f(\tau) \tag{2.9}$$

$$\frac{d^2}{dp^2}\left((1 - p^2)p_e(p)\varphi(p)\right) - \frac{d}{dp}\left(I(u - p)p_e(p)\varphi(p)\right) = -\lambda p_e(p)\varphi(p) \tag{2.10}$$

Assuming discrete eigenvalues, Eqn. 2.9 can be easily solved to yield,

$$f_n(\tau) = k_n e^{-\lambda_n \tau} \tag{2.11}$$

while using Eqn. 2.8 in Eqn. 2.10 gives the Sturm-Liouville form,

$$\frac{d}{dp}\left((1 - p^2)p_e(p)\frac{d\varphi(p)}{dp}\right) + \lambda p_e(p)\varphi(p) = 0 \tag{2.12}$$

Necessary and sufficient conditions for Eqn. 2.12 to yield a complete orthonormal set of polynomials as eigenfunctions have been studied by Wong et. al. [69]. They can be summarized as follows:

$$B(p_1)p_e(p_1) = B(p_2)p_e(p_2) = 0, \tag{2.13}$$

$$\text{where} \quad p_1 \leq p \leq p_2$$

$$A(p) = ap + b \tag{2.14}$$

$$B(p) = cp^2 + dp + e \quad \text{and} \tag{2.15}$$

$$\int_{p_1}^{p_2} p^n p_e(p) dp < \infty, \quad n = 0, 1, ..., n < \infty \tag{2.16}$$

From Eqn. 2.5,2.6 and noting that $-1 \leq p \leq 1$, it is easy to see that the necessary and sufficient conditions are satisfied. The above conditions restrict the density function $p_e(p)$ to be of the form [69],

$$p_e(p) = \frac{1}{2^{\alpha+\beta+1}} \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\Gamma(\beta+1)} (1-p)^{\alpha}(1+p)^{\beta}, \quad \alpha, \beta > -1 \tag{2.17}$$

while the polynomial eigenfunctions $\varphi_n(p)$ orthonormalized with respect to the equilibrium density function $p_e(p)$ are the Jacobi polynomials,

$$\begin{aligned} \varphi_n(p) = & \frac{(-1)^n}{2^n} \times \sqrt{\frac{(2n+\alpha+\beta+1)\Gamma(n+\alpha+\beta+1)}{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)}} \\ & \times \sqrt{\frac{\Gamma(\alpha+1)\Gamma(\beta+1)}{\Gamma(\alpha+\beta+2)n!}} \times (1-p)^{-\alpha}(1+p)^{-\beta} \\ & \times \frac{d^n}{dp^n}\left[(1-p)^{n+\alpha}(1+p)^{n+\beta}\right] \end{aligned} \tag{2.18}$$

For $p_e(p)$ defined as in Eqn. 2.17, the functions

$$\begin{aligned} A(p) &= \gamma(\beta - \alpha) - \gamma(\alpha + \beta + 2)p \\ &= Iu - Ip \qquad \text{(from Eqn.2.6)} \\ B(p) &= 2\gamma(1 - p^2) \\ &= 2(1 - p^2) \qquad \text{(from Eqn.2.5)} \\ \text{and} \quad \lambda_n &= \gamma n(n + \alpha + \beta + 1) \end{aligned} \tag{2.19}$$

Solving 2.19 yields

$$\begin{aligned} \gamma &= 1, \\ \lambda_n &= n(n + I - 1), \\ \alpha &= I_L - 1 \text{ and} \\ \beta &= I_R - 1 \end{aligned} \tag{2.20}$$

This restricts $I_R, I_L \geq 1$. The joint probability density function $p(p_0, p; \tau)$ have the form,

$$p(p_0, p; \tau) = p_e(p_0)p_e(p) \sum_{n=0}^{\infty} e^{-\lambda_n \tau} \varphi_n(p_0)\varphi_n(p) \tag{2.21}$$

where p_e, φ_n and λ_n are given by respectively Eqns. 2.17, 2.18 and 2.19, and initial polarization factor $p_0 = p(\tau_0)$. This completely specifies the progression of the joint probability density function.

The results shown in Fig. 2.2 are for $N = 200$, $p_0 = 0$, $I_R = 11$ and $I_L = 2$. In other words, initially exactly 50% of the 200 undecided evacuees start moving right and 50% start towards the left exit. As they start moving as two discrete groups, there is opinion exchange and a few people change their mind and join the other group moving in the opposite direction. Figure 2.2 shows the probability distribution of how the crowd is expected to be polarized at each subsequent time steps. Numerical results from 2500 Monte Carlo simulations overlayed on the analytical results verify the accuracy of the results and the validity of the assumptions made. Interestingly, with increasing number of average interactions per person, the probability distribution flattens out, while the mean slowly moves towards higher values of p. The gradual favoring of the right exit by more people is

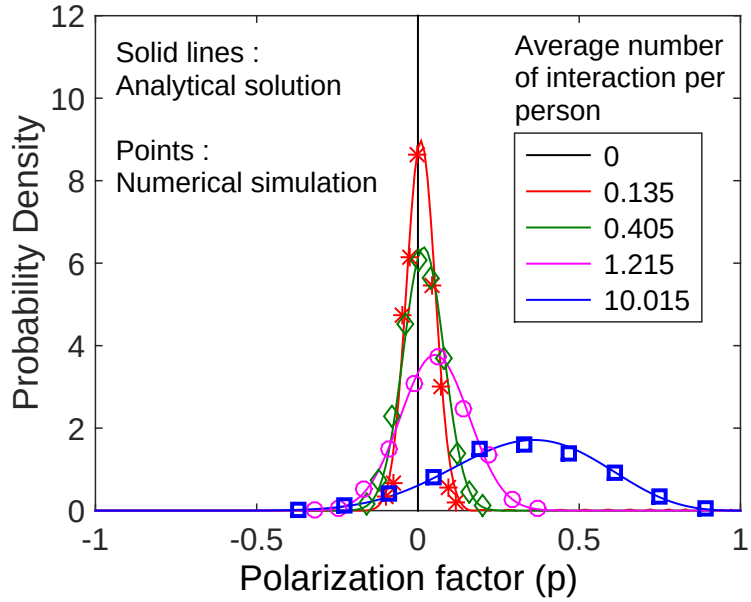


Figure 2.2: Analytical and numerical results for probability distribution of final crowd polarization factor with different number of average interaction per person. Here, $N = 200$, $p_0 = 0$, $I_R = 11$ and $I_L = 2$.

a result of the larger number of independent nodes moving to the right ($I_R = 11$) compared to the left ($I_L = 2$).

From the point of view of faster evacuation, it is beneficial to be able to influence the final polarization to match the flow capacity of the individual exits. For example, in our experiments, if the right exit has twice the flow capacity of the left exit, then it is preferred that the crowd polarization (p) is equal to 0.33. The analytical and numerical results suggest that the the presence of strong opinion holders has an enormous effect on polarizing the crowd, thereby affecting the total evacuation time by utilizing the available exits more or less effectively.

2.2.2 Constant velocity dynamics

In the previous section, movement of individuals from one group to another is assumed to occur at a faster time scale compared to the group movement. This dynamics was modeled on the assumption that individuals move faster than a tightly packed crowd trying to navigate a narrow corridor. But this assumption fails to hold if we consider a larger space where individuals are free to move at their own pace, limited only by their physical capabilities. In

that scenario, formation of distinct clusters of people moving together is unlikely, rather a more uniformly spread out distribution over the movement axis seems to be more probable. To investigate the implications of this scenario, a constant velocity model is investigated next. Each node, i is assumed to be moving at their maximum speed toward their respective choice of exit, σ_i, where $\sigma_i \in \{E_L, E_R\}, \forall i \in N \cup I$. Unlike the previous case, the strength of interaction between two individuals is now modeled as a function of the distance separating them. In this case, we model the strength of interaction as $SOI(d_{ij}) = e^{-\delta \times d_{ij}}$, where δ is the decay rate and d_{ij} is the distance between nodes i and j at that instant. Incorporating the SOI factor, the modified Voter model dynamics is now as follows (Alg. 1).

Algorithm 1: Hybrid motion model with strength of influence voter model

```
Data: N, I_R, I_L, δ
Result: Decision sharing model with SOI
Initialization: p_0, {σ_i : i ∈ N ∪ I};
while egress is not complete do
    while each node hasn't interacted once do
        Select each node i in random order, where i ∈ N;
        Select random neighbor j for each, where j ∈ N ∪ I ;
        Determine SOI(d_ij) = e^(−δ×d_ij);
        Set σ_i = σ_j with probability SOI(d_ij);
    end
    Each node i moves one step towards their exit choice σ_i, where i ∈ N ∪ I
end
```

Essentially, for a higher decay rate, the interaction is similar to that implemented with the previously discussed narrow central interaction zone, inside which all interactions are constrained to occur. For lower decay rates, even more distant individuals have a higher probability of successfully changing the opinion of the other. This strength of interaction creates a personal interaction zone for each individual separately and it moves with the individual. The size of the interaction zone is determined by the decay rate (2.3).

Leaders are recognized by their ability to influence a large number of people. This is modeled by relaxing the distance restriction on the SOI for such individuals, i.e., leaders are assumed to be able to influence undecided individuals successfully, regardless of the distance between them. With this setup various numerical simulation experiments were carried out and the results are presented and discussed in the following section.

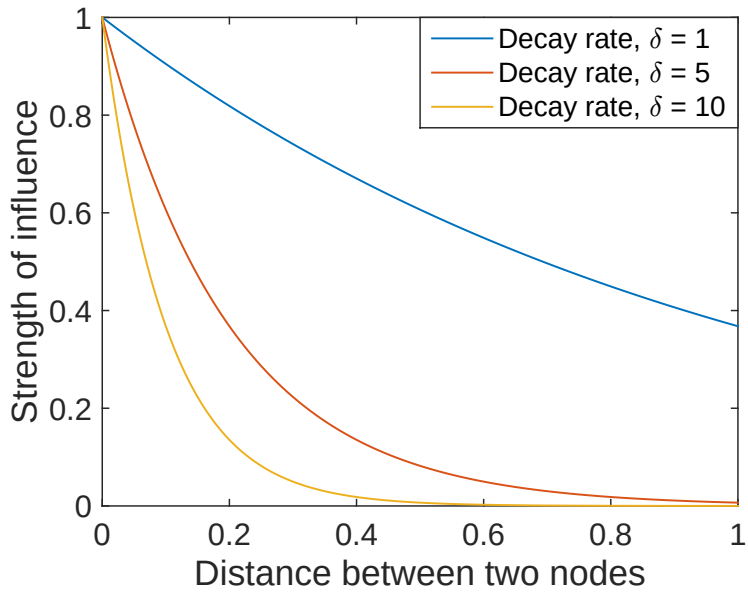


Figure 2.3: Strength of influence between two nodes with different decay rates (δ)

2.3 Results and Discussion

2.3.1 Movement without leaders ($I = 0$)

The first set of simulations tries to isolate the effect of initial bias and the effect of varying degree of interactions between evacuees. All simulations were conducted with $N = 100$ without the presence of any strongly opinionated individuals (i.e. $I = 0$). The distribution of final crowd polarization were obtained by running identical experiments 2500 times. The top row in Fig. 2.4 shows the final distribution of crowd polarization (p) with increasing δ and $p_0 = 0$. The bottom row is for $p_0 = 0.5$, to show the effect of starting with a relatively higher initial polarization. We can interpret that with more interactions amongst the individuals, they end up coalescing completely at either one of the exits (Fig. 2.4a). If the initial crowd polarization is non zero ($p_0 \neq 0$) then the crowd coalesce more at the exit towards which they are initially biased (Fig. 2.4d). When the decay rate (δ) is increased, the number of successful interactions goes down and hence the distribution of crowd at the exit become less predictable. The crowd does not get enough chances to successfully interact and coalesce to a unified decision before they reach the exits (Fig. 2.4c). The initial crowd bias helps to tilt the final distribution towards the respective exit nevertheless. The entire crowd ending up

in one of the two exits is generally undesirable unless the state of emergency renders one of the exits unusable.

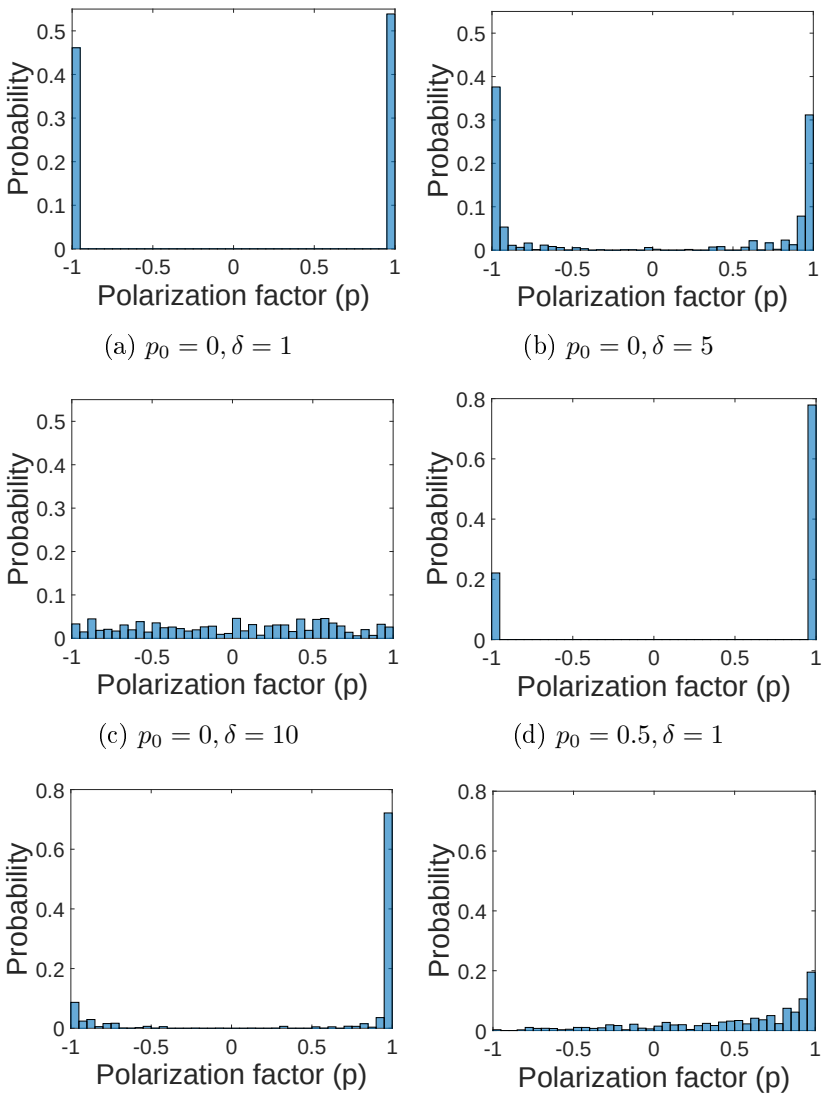
Figure 2.5(c) shows the plot of final polarization factor characteristics (mean and entropy) with different strength of interactions. The mean of final polarization factor (in the absence of leaders) depends only on the initial polarization (p_0), but independent of the amount of interactions among nodes. This reinforces the previous argument that the initial crowd bias helps to tilt the final distribution towards the corresponding exit. The entropy is low for lower δ. This conveys that with more interaction the final distribution become more ordered. The entropy goes up with higher δ since the distribution become less predictable. With more initial crowd bias the entropy goes down as the $p_0 \neq 0$ creates a more ordered initial crowd opinion leading to a relative more ordered final crowd opinion.

Figures 2.5(a) and 2.5(b) show the plots of mean location of the groups moving respectively towards right and left. With lesser interaction the crowd moves quickly towards their respective exit. This is expected since with more interactions among the individuals there is more possibility for them to switch their exit choice midway and thus end up increasing the average number of steps required to reach their desired exit. With a initial biased population towards the right exit ($p_0 > 0$), the average number of steps required by the crowd moving towards the right exit decreases and the average number of steps required by the crowd moving towards left exit increases. Since the initial bias of the crowd reinforces the right opinionated group and conflicts with the left opinionated group, the movement towards the exit in the right side is bolstered and the movement in the opposite direction is impeded. The next sub-section delves into the dynamics of the crowd in the presence of strong opinion holders ($I > 0$).

2.3.2 Movement in the presence of leaders ($I > 0$)

Constant u - Variable I

The next set of experiments were conducted to study the effect of global influence ratio (ζ) during egress. The influencer ratio, i.e $u = (I_R - I_L)/I = -1/11$ is kept constant; initial polarization is maintained at $p_0 = 0$. As in previous section, $N = 100$. Figure 2.6

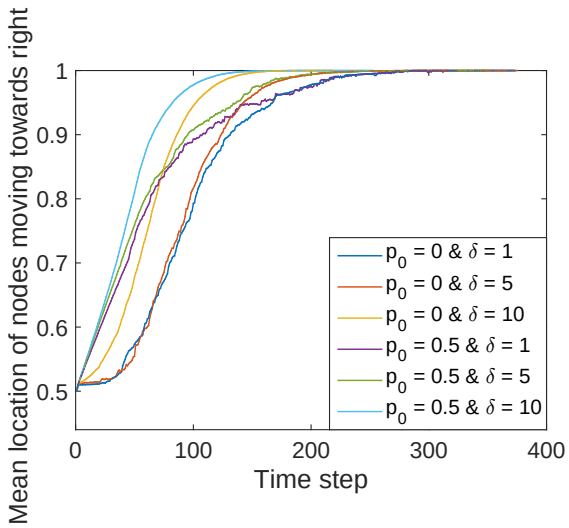


(a) $p_0 = 0, \delta = 1$ (b) $p_0 = 0, \delta = 5$

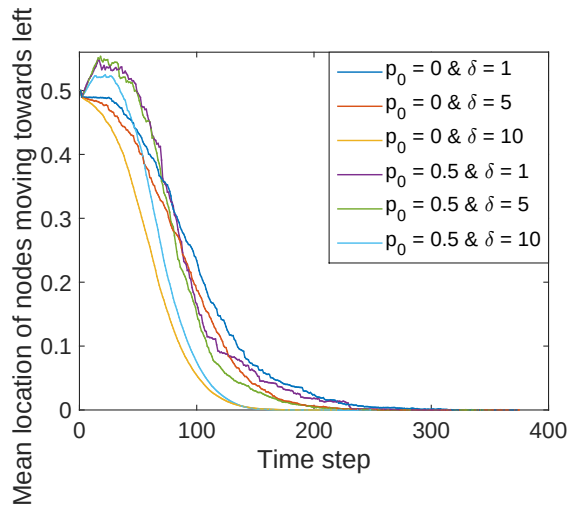
(c) $p_0 = 0, \delta = 10$ (d) $p_0 = 0.5, \delta = 1$

(e) $p_0 = 0.5, \delta = 5$ (f) $p_0 = 0.5, \delta = 10$

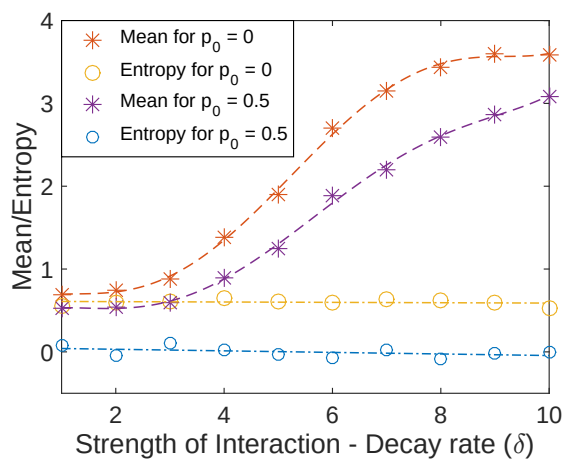
Figure 2.4: The effect of p_0 and δ on the final distribution (at the exit) of polarization factor p. Here, $N_R + N_L = 100$ and $I = 0$.



(a)



(b)



(c)

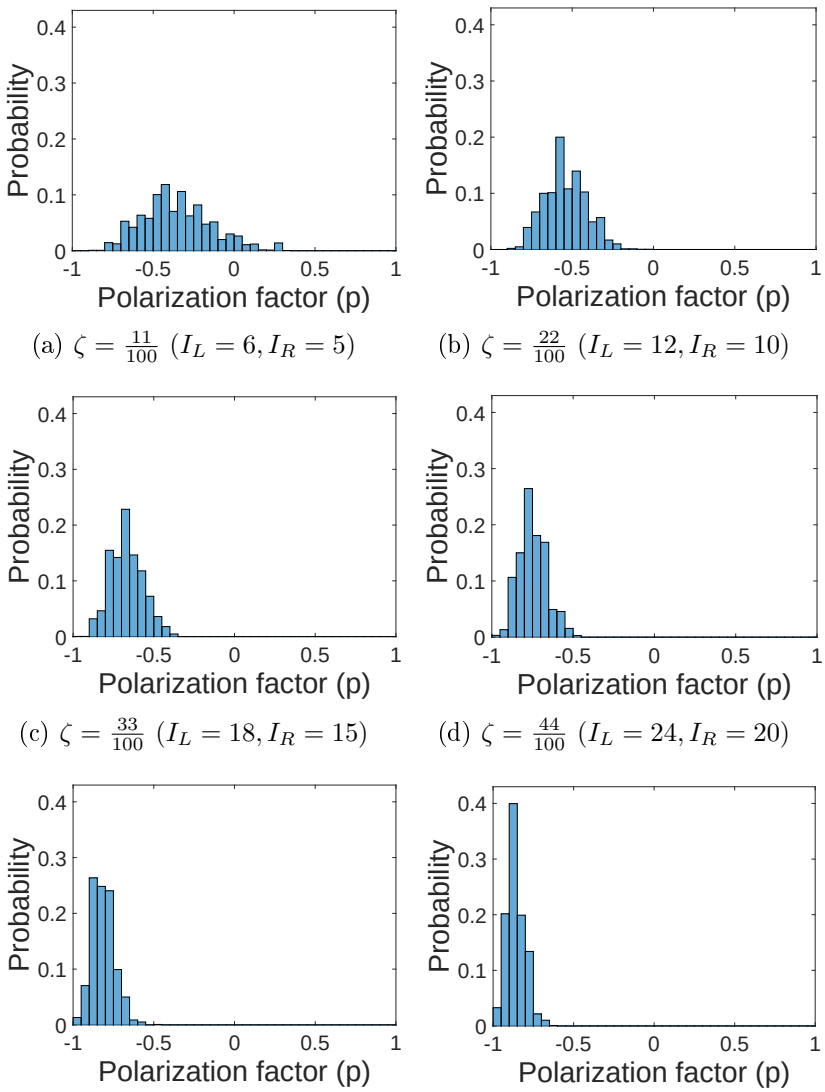
Figure 2.5: (a) Movement dynamics for nodes moving towards right exit, (b) Movement dynamics for nodes moving towards left exit and (c) Final polarization factor characteristics with different decay rates and initial polarization factors ($p_0 = 0$ and $p_0 = 0.5$). For all graphs $N_R + N_L = 100$ and $I = 0$.

shows the distribution of the final polarization of the undecided crowd when varying number of influencing nodes are embedded in the crowd. The distribution was obtained through running the experiment under the same conditions 2500 times. The graphs point out two significant characteristics. With greater magnitude of I, the distribution of polarization factor becomes sharper and shifts towards the side with more number of influencers, in this case towards the left since $I_L > I_R$. With an increasing global influence ratio, ζ, their reach expands and thus they are able to impact the final outcome with more certainty.

Figure 2.7(c) displays the mean and entropy of the equilibrium p distribution with varying ζ. The mean shifts towards the side with higher number of influencers and the entropy decreases as the distribution becomes sharper. With more influencers in the crowd, the probability of successful interaction increases since the influencers are not restricted by the distance rule and thus brings down the entropy, i.e. uncertainty in the outcome. The movement dynamics of the crowd is depicted in Figs. 2.7(a) and 2.7(b). Since the crowd is attracted to move towards the left exit by a larger number of strongly opinionated individuals, the movement towards the left is quicker compared to the movement towards the opposite side. But, there is a detrimental effect with increasing number of leaders. The average number of steps required by the crowd to reach an exit goes up and this is the effect of a larger number of successful interactions which implies that individuals are more likely to remain indecisive and thus they end up in the corridor for longer period. The next set of experiments were modeled to study the effect of influencer ratio u on the crowd dynamics.

Constant I - Variable u

The distribution of the crowd polarization at the exit with $N = 100$, initial condition $p_0 = 0$, $\delta = 10$ and $I = 10$ with different u is illustrated in Fig. 2.8. As in previous sections, the distribution was obtained by running the simulation 2500 times under same initial conditions. The more skewed the influencer ratio, the higher the probability that the crowd moves en masse towards that particular exit. Even, the presence of strongly opinionated individuals evenly attracting towards both exit, (i.e. $u = 0$) has a desirable effect on the crowd dynamics. The distribution is more condensed than in the case with no influence at all ($I = 0$). Figure 2.9(c) presents the mean and entropy of the final polarization factor for different u. The

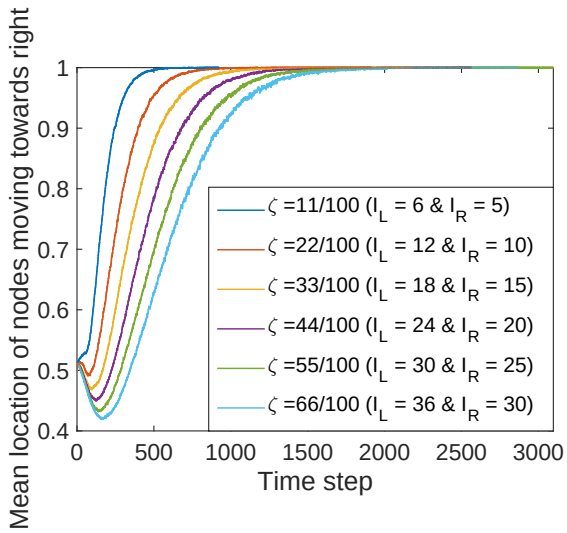


(a) $\zeta = \frac{11}{100}$ $(I_L = 6, I_R = 5)$ (b) $\zeta = \frac{22}{100}$ $(I_L = 12, I_R = 10)$

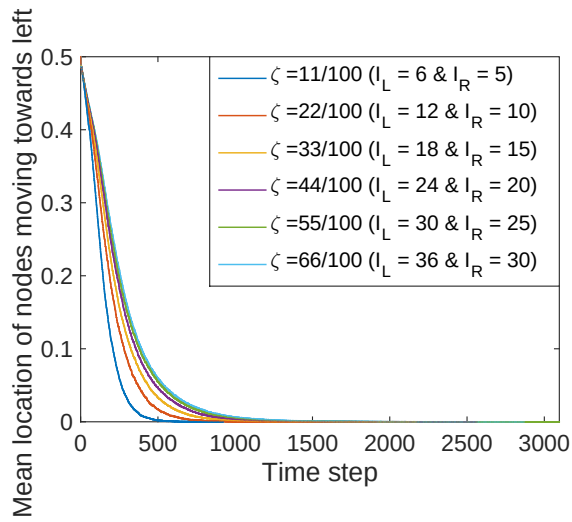
(c) $\zeta = \frac{33}{100}$ $(I_L = 18, I_R = 15)$ (d) $\zeta = \frac{44}{100}$ $(I_L = 24, I_R = 20)$

(e) $\zeta = \frac{55}{100}$ $(I_L = 30, I_R = 25)$ (f) $\zeta = \frac{66}{100}$ $(I_L = 36, I_R = 30)$

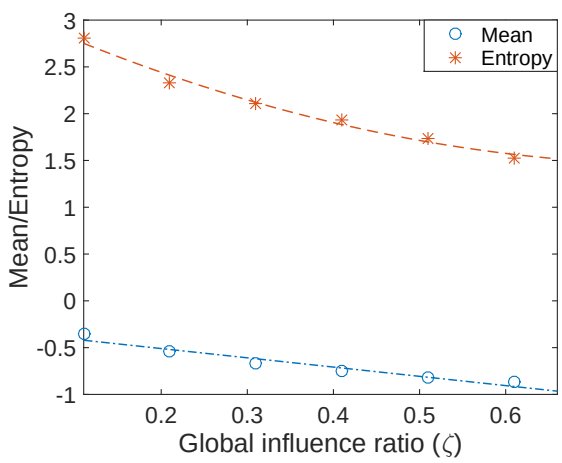
Figure 2.6: Effect of global influence ratio (ζ) with initial polarization $p_0 = 0$, $u = \frac{-1}{11}$, $\delta = 10$ and $N_R + N_L = 100$



(a)

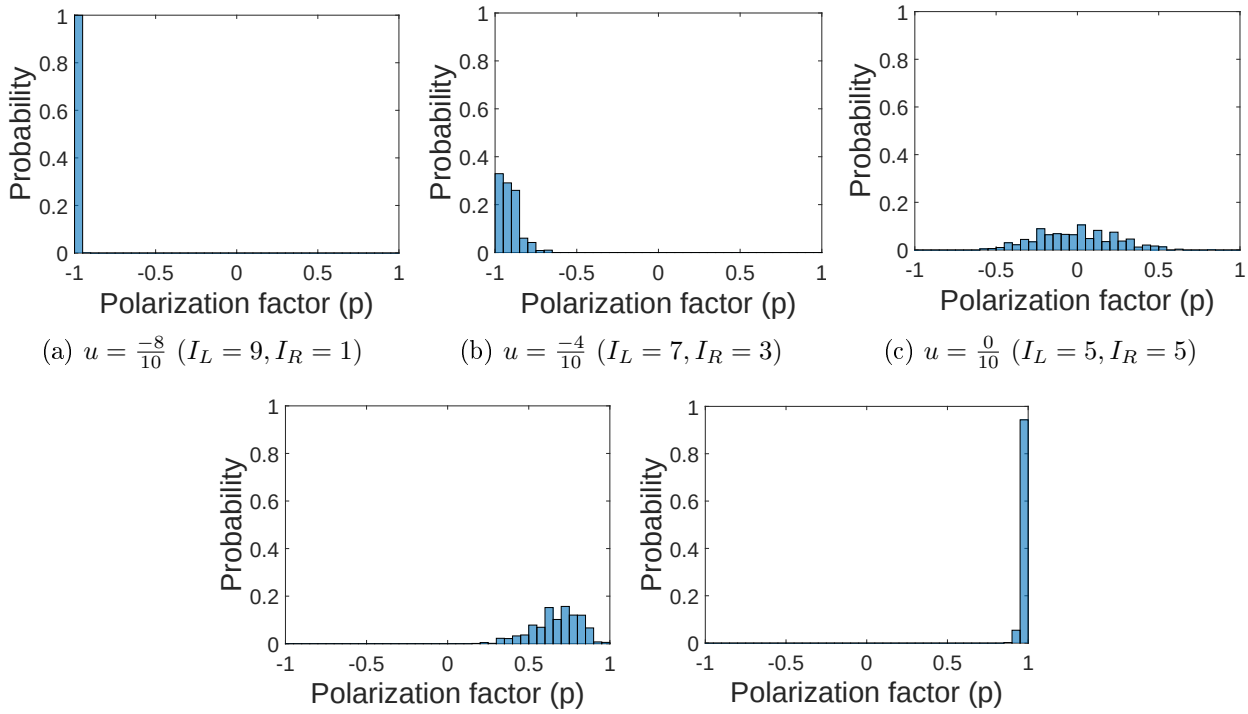


(b)



(c)

Figure 2.7: (a) Movement dynamics for nodes moving towards right exit, (b) Movement dynamics for nodes moving towards left exit and (c) Final crowd polarization characteristics for different ζ. For all graphs $u = \frac{-1}{11}$, $p_0 = 0$, $\delta = 10$ and $N_R + N_L = 100$.



(a) $u = \frac{-8}{10}$ $(I_L = 9, I_R = 1)$ (b) $u = \frac{-4}{10}$ $(I_L = 7, I_R = 3)$ (c) $u = \frac{0}{10}$ $(I_L = 5, I_R = 5)$

(d) $u = \frac{2}{10}$ $(I_L = 4, I_R = 6)$ (e) $u = \frac{6}{10}$ $(I_L = 2, I_R = 8)$

Figure 2.8: Effect of different influencer ratio (u) with initial polarization $p_0 = 0$, $I = 10$, $\delta = 10$ and $N_R + N_L = 100$

mean has monotonic but non-linear correlation with the influencer ratio (u). The entropy falls as $abs(u)$ increases, since the distributional uncertainty is reduced the more skewed the influence on the population. The influencers ensure that the crowd coalesce more predictably with $u \neq 0$. Figure 2.9(a) depicts the movement dynamics of the crowd moving toward the exit on the right side for $u > 0$. With increasing u from 0 to 1 the movement towards the right side exit becomes quicker since the influencers attract the crowd towards the right side exit more strongly. Figure 2.9(b) portrays the movement of crowd which movers towards the exit on the left side of the corridor for $u < 0$. With decreasing u from 0 to -1, the average number of steps required by the crowd to egress through the left side exit goes down. The influencers are able to shepherd the crowd more effectively towards the left side exit with decreasing u. Thus it can be concluded that with lesser total number of strong opinion holder (I) and $u \neq 0$, the crowd can be split into any ratio for optimally utilizing the exits and thus achieve quicker evacuation of the crowd from the hazardous situation.

Constant u and I - Variable p_0 and δ

The last set of experiments were conducted to study the effect of different initial bias (p_0) and decay rate of communication (δ) with constant numbers of strongly opinionated individuals ($I_L = 5$ and $I_R = 2$) amongst the crowd ($N = 100$). Figure 2.10(a) brings out the characteristics of final crowd polarization factor with different initial crowd polarization (p_0) and decay rates (δ). With a small number of strong opinion holders, the mean of the final polarization factor is only slightly affected by the initial crowd bias for different strength of interaction and different initial crowd polarization. This leads to the conclusion that with a relatively few strong opinion holders the crowd can be directed such that they end up utilizing the exits optimally.

Figure 2.10(b) shows the effect of strength of interaction on the final polarization factor characteristics. With lesser interactions, the effect of strong opinion holders on the mean diminishes slightly. This is because individuals other than the influencers have lesser probability of successful interactions and thus the secondary passing of influencers' opinions is restricted with increasing δ. From an information content point of view, the entropy decreases when the initial crowd bias (p_0) favors the influencer ratio ($p_0 < 0$ and $u < 0$

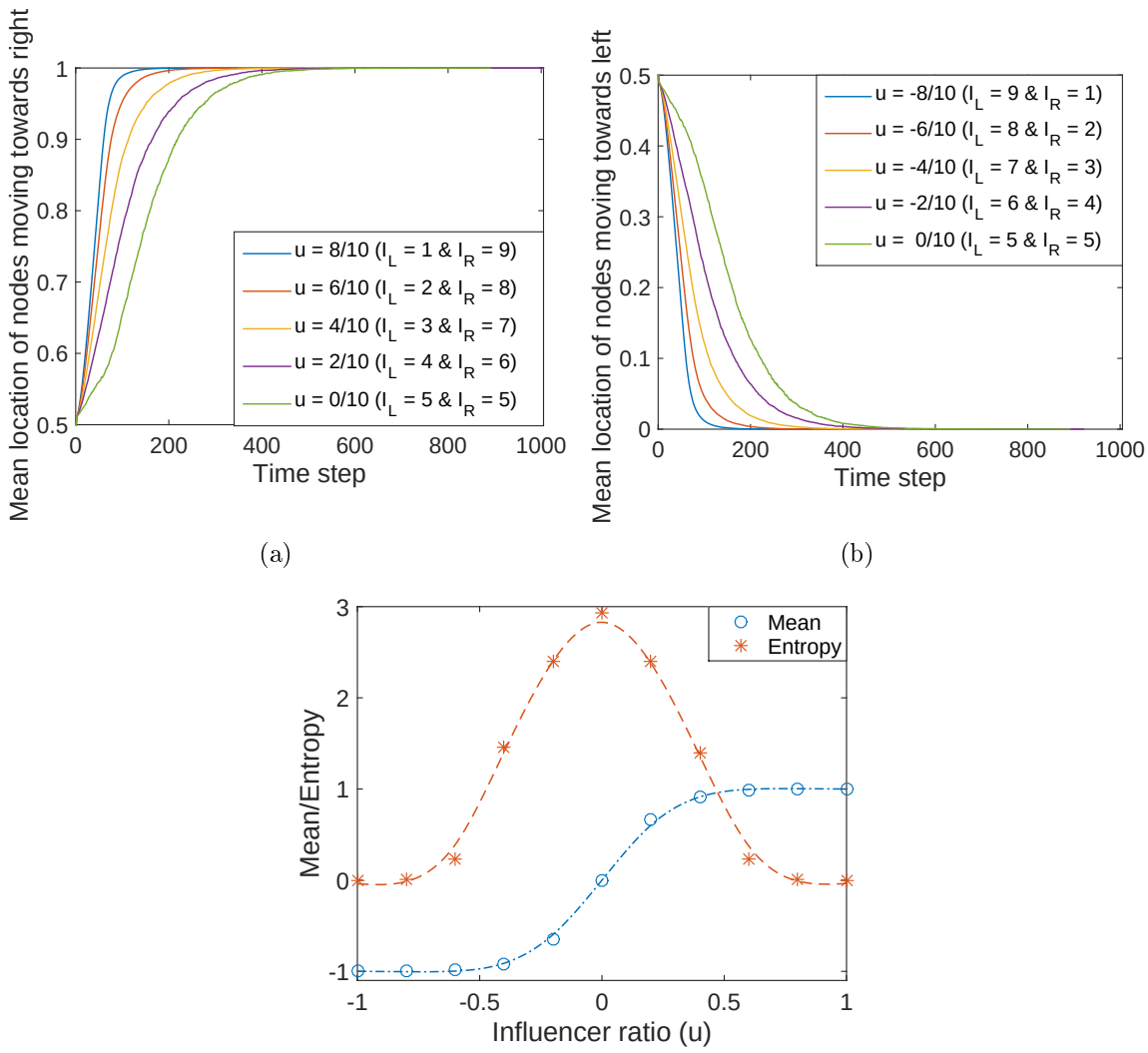
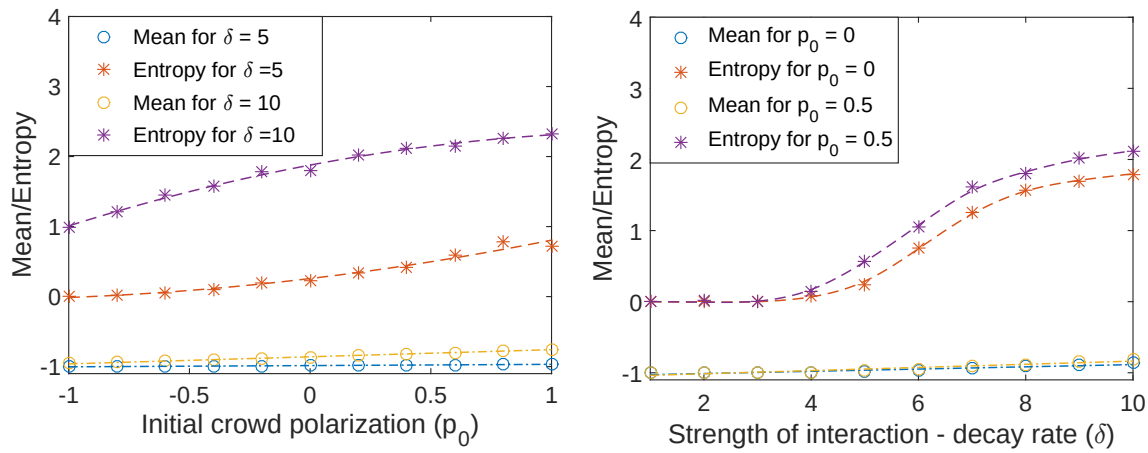


Figure 2.9: (a) Movement dynamics for nodes moving towards right exit, (b) Movement dynamics for nodes moving towards left exit and (c) Final crowd polarization characteristics for different u. Here, $p_0 = 0$, $I = 10, \delta = 10$ and $N_R + N_L = 100$.



(a) (b)

Figure 2.10: Final crowd polarization characteristics (a) For different p_0 and (b) For different δ. $N_R + N_L = 100$.

or $p_0 > 0$ and $u > 0$). Since the initial crowd polarization and influencer ratios reinforce one another the uncertainty and consequently the entropy goes down. When the initial crowd polarization opposes the influencer ratio ($p_0 < 0$ and $u > 0$ or $p_0 > 0$ and $u < 0$), the entropy increases. The entropy increases with increasing δ. With lesser probability of successful interaction, the effect of influencer propagate more slowly and hence the increase in entropy with increasing δ.

2.4 Summary

This part of the research is unique in the sense it combines a motion model with a explicit opinion sharing model to study the effects of opinion sharing on crowd evacuation from a long corridor with exits at each end. People with leadership skills and strong bias towards a particular exit play a pivotal role in determining how the crowd is dynamically attracted towards each of the exits. The effect of leaders on the dynamics of the hybrid model is studied in detail.

In contrast to existing models, which usually focuses more on developing realistic motion models, this work tries to combine the effect of opinion sharing and movement among egressing individuals and also discusses interesting effects of strongly opinionated leaders

in shaping the crowd movement dynamics. Additionally, different strengths of interaction were tested and an analytical solution for interactions within a restricted zone were presented.

Chapter 3

Parametric study of egress dynamics in a Markov Decision Process framework with spatially bounded opinion sharing

3.1 Objective

A common theme among the reviewed literature was that there is a need for a parametric study of a egress model which incorporates both movement and decision with an explicit sharing and mimicking of decisions among the evacuees. The previous chapter presented an analytical and simulation result for explicit opinion sharing during evacuation through a long corridor [70]. The movement model was kept simple and the main focus was on opinion propagation through the evacuating crowd. The decision model was based on binary choice (i.e.,) picking either one of the two available exits. In this work we introduce a more naturalistic movement model which incorporates collision avoidance with pedestrians and walls. The personal space of the pedestrians were not violated. The decision model takes into account impatience wherein a pedestrian may become impatient with their current choice of exit due to bottlenecking/crowding and switch to a different route. A spatial boundary which mimics the visual range of an individual is utilized and people within the boundary

affect the opinion of the individual. This account for the herding behavior [71] prevalent in existing literature. The simulation model is presented in detail in the following section*.

3.2 Simulation setup

3.2.1 Movement model

To study the egress dynamics with decision framework of individuals and explicit opinion sharing, the building setup was designed as shown in Fig. 3.1. The building consists of two rooms which open into a corridor. A person can choose to move towards either end of the corridor. At either end of the corridor they will decide between the two final exit points. The rooms are populated with people from different age and gender groups and given walking speed accordingly. Each individual was assumed to occupy a circle of $1ft$ radius with an additional $1ft$ radius designated as personal space. Let the total number of people in the building be denoted by N, the velocity of the individuals by V_i, the current position of the individuals by (x_i^t, y_i^t) and the individuals' desired exit point by E_i (provided by the decision model). The movement model for each individual is given by Alg. 2.

Thus every individual attempts to move at every time instant (every 1 *sec*) respecting others' personal space and avoiding collision with walls and people. The underlying decision and opinion sharing model are explained in subsequent sections.

3.2.2 Decision model

The underlying decision logic for individuals is modeled as a Markov decision process ([72, 73]). A Markov decision process is defined by $M = \{S, A, P, \gamma, R\}$ where:

S is the set of all possible decision states,

A is the set of all available decision/actions,

*Results presented in this section are partially from our accepted work: Srinivasan, A. R., Karan, F.S.N., & Chakraborty, S. (2018, July).A study of how opinion sharing affects emergency evacuation. In International Conference on Social Computing, Behavioral-Cultural Modeling & Prediction and Behavior Representation in Modeling and Simulation. Springer, Cham

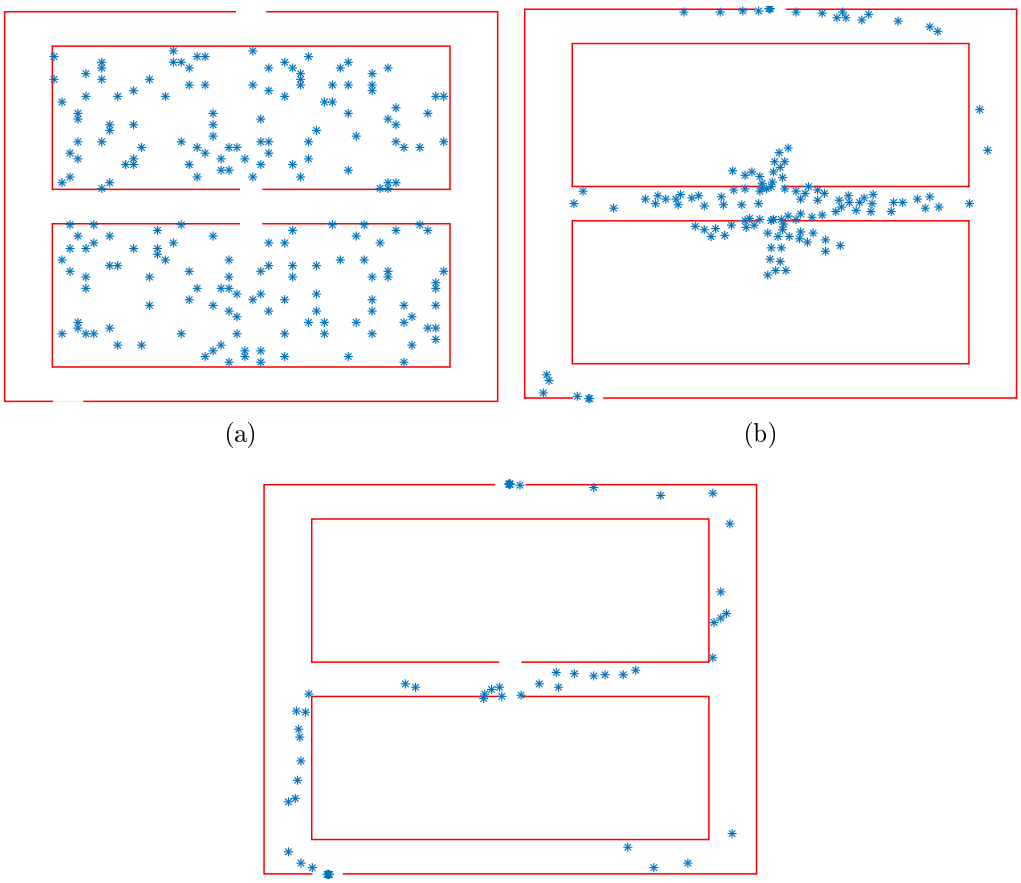


Figure 3.1: (a)Snapshot of 200 people at the start of an egress ($t = 1$ *sec*), (b) Snapshot of the people in the middle of an egress ($t = 100$ *sec*), and (c) Snapshot of people near the end of an egress ($t = 200$ *sec*)

Algorithm 2: Movement model for each individual

Data: $N, V_i, (x_i^t, y_i^t), E_i$
Result: (x_i^{t+1}, y_i^{t+1})

```
while Individual hasn't exited the building do
    Convert the cartesian coordinates (x_i^t, y_i^t) to polar coordinates (r_i^t, θ_i^t)
    Shift the origin to desired exit point E_i (midpoint of the exit)
    Obtain new position by computing r_i^{t+1} = r_i^t − V_i and de-shift the origin and
      convert to cartesian coordinates (x_i^{t+1}, y_i^{t+1})
    if new position (x_i^{t+1}, y_i^{t+1}) is within anyone else personal space then
        Stay at old position (x_i^{t+1}, y_i^{t+1})=(x_i^t, y_i^t)*
    end
    if collision with walls then
        Reduce walking speed V_i till no wall collision
    end
end
* They try for the farther corners of the exit before staying at old position
```

P is the transition probability $P(s, a, s')$. It gives the probability an individual assigns for successful physical transition to state s' from state s after deciding to take action a,

R is the set of rewards - This indicates the mental payoff assigned to the various decision states by an individual. The individual's overall route choice depends on the reward structure,

γ is the discount factor $\in [0, 1)$ - This is used to make the computation of accumulated rewards mathematically tractable.

Each individual has exits 1, 2, 3, 4, and 5 marked as E_1, E_2, E_3, E_4, E_5, and the trails connecting the exits, marked as T_{ij} (see Fig. 3.2) as available decision states. T_{ij} denotes the corridor connecting the ith exit to the jth exit. Every individual can decide to move towards one of the immediately available exit points and they will land in the state corresponding to their current position. For example, if a person in the corridor outside the room decides to move towards exit 2, his/her state would be T_{12}. When the same person is physically near exit 2 they can utilize the exiting action e and move to state E_2. Therefore, the set of available actions consist of decisions to move towards exits and the action of exiting labeled as e_1, e_2, e_3, e_4, e_5, and e respectively in Fig. 3.2.

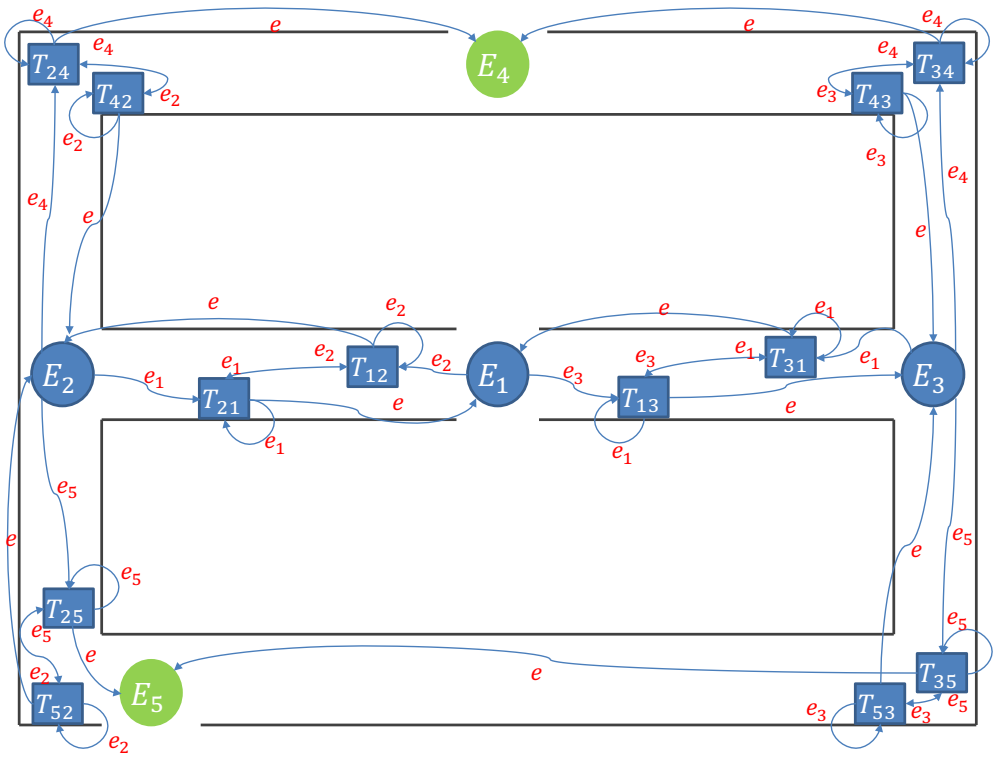


Figure 3.2: Illustration of the egress setup with underlying decision model

Initially, the transition probability ($P(s, a, s')$) for all state and action pairs is set at 0.9. To complete the transition probability definition, $P(s, a, s) = 1 - P(s, a, s')$. This takes into account the environmental uncertainties. The transition probability for action e ($P(s, e, s')$) is modified as time progresses to account for the physical reality. An individual estimates his/her travel time to the desired exit point when they start their egress towards that particular exit point. If the individual hasn't reached the desired position in their estimated time they start to get impatient. Correspondingly, the chance of success they had assigned for that particular state transition starts to decay exponentially as expressed by,

$$P(s, e, s') = P(s, e, s') * exp(-\alpha \times t_{diff}) \tag{3.1}$$

where t_{diff} =Time spent in state T_{ij} − Estimated travel time to exit E_j.

We have experimented with 3 different impatience growth rate, α to simulate different crowd behaviors. The transition probability decay with different impatience rates are

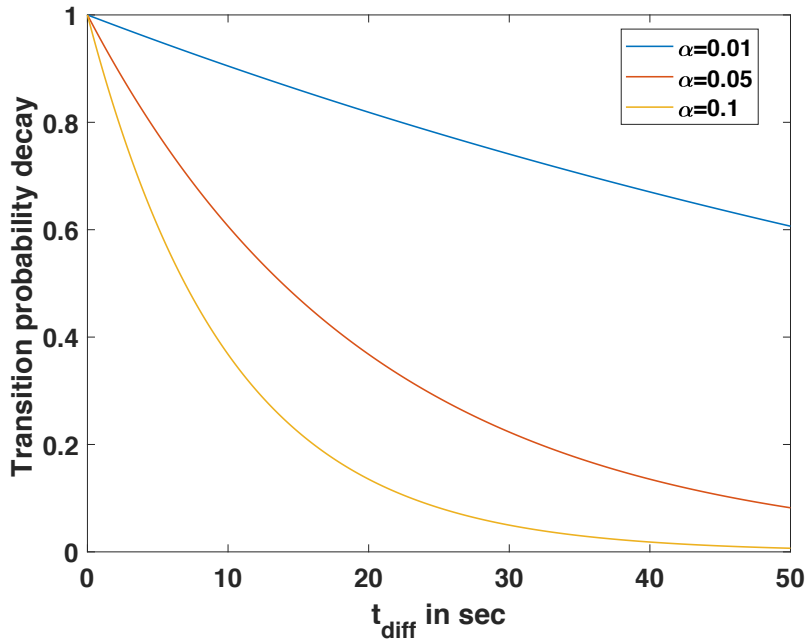


Figure 3.3: Illustration of the transition probability decay with different impatience rate

illustrated in Fig 3.3. With $\alpha = 0.1$ we can simulate an highly impatient crowd whereas with $\alpha = 0.01$ we can approximate a crowd that is relatively more patient.

The exits from the building (exits 4 and 5) are given the maximum reward magnitude and the immediate exit before reaching them (exit 2 or 3) are given lower reward magnitude and the exit from the rooms (exit 1) is given even lesser reward magnitude. The trail states are given rewards that are inversely proportional to the trail length and the maximum reward for the trail is upper bounded by the minimum reward for all the exits. The building setup was designed with one obvious shortest path, a couple of paths of moderate length and a longer path for safe evacuation. Individuals will typically chose the shortest path. However, if the lanes are crowded, then they tend to move towards the next best available route. The goal of each decision maker is to reach either exit 4 or 5 as quickly as possible. Physically it means they have successfully exited the building. Verlander and Heydecker [74] reported, based on an empirical study, that pedestrian prefer shortest route. This reward structure enables the decision maker to seek the decision state that leads to the shortest path towards the exit, but the framework allows individuals to change their decision if the are unable to reach their desired exit within a reasonable time frame.

Individuals are assigned a decision timer (τ_i) from a normal random distribution. Each individual performs a planning routine whenever their decision timer expires. For planning their route, individuals compute the value of available states (exits and trails), compare the values, and decide to move along the trail with the highest value. The value of a state is the expected cumulative reward that can be obtained from that state. The discount factor is used in the summation to weigh the immediate reward more than the future rewards. Formally, a value iteration algorithm ([72, 73]) is used to find the value of states and it is given in Alg. 3 .

Algorithm 3: Value iteration algorithm

Data: $M = \{S, A, P, \gamma, R\}$
Result: $V(s)\ \forall\ s \in S$
Initialize: $V_0(s) \leftarrow 0, \forall s \in S$
repeat
  $V_{i+1}(s) = \max_a \sum_{s' \in S} P(s, a, s')[R(s') + \gamma V_i(s')]$
until $\max_s |V_{i+1}(s) - V_i(s)| < \epsilon$;

The value of states found with value iteration algorithm satisfies the Bellman optimality condition. The Bellman optimality condition states that the action taken at a state has to result in landing at the best possible next state with respect to their calculated value. Thus each individual optimizes his/her route at every decision cycle.

3.2.3 Spatially bounded confidence model

Humans have a tendency to herd ([75]) and it is captured in this paper with a spatially bounded confidence model. Previous studies have mostly concentrated on mathematical modeling of just the opinion space [76, 77, 78, 79].The bounded confidence model has been utilized to model opinion sharing in [80, 81, 82, 83]. Opinion is conceived as a continuous quantity and nodes with similar opinion (i.e., within a confidence boundary) interact with each other and change their opinion state. This model is modified to suit the egress dynamics by using distance between individuals as the confidence boundary metric. Each individual after completing a value iteration cycle will interact with individuals within their herding

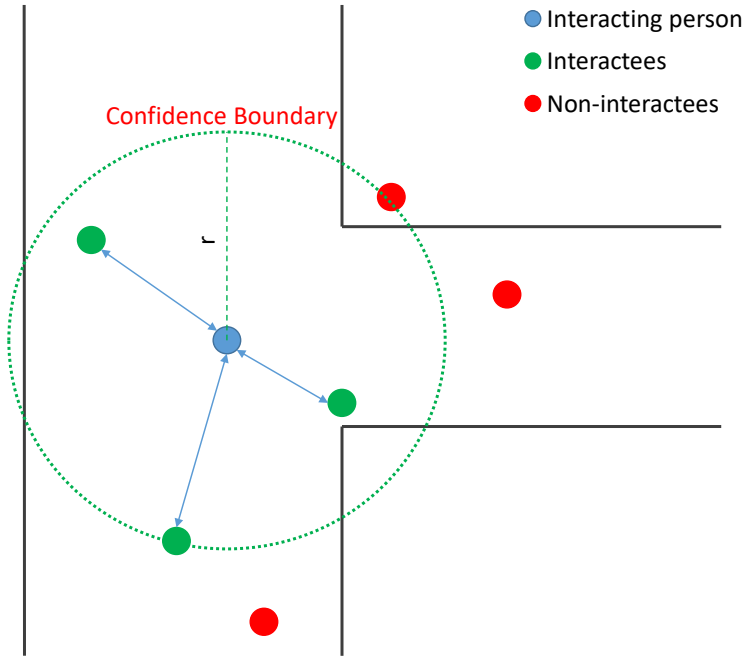


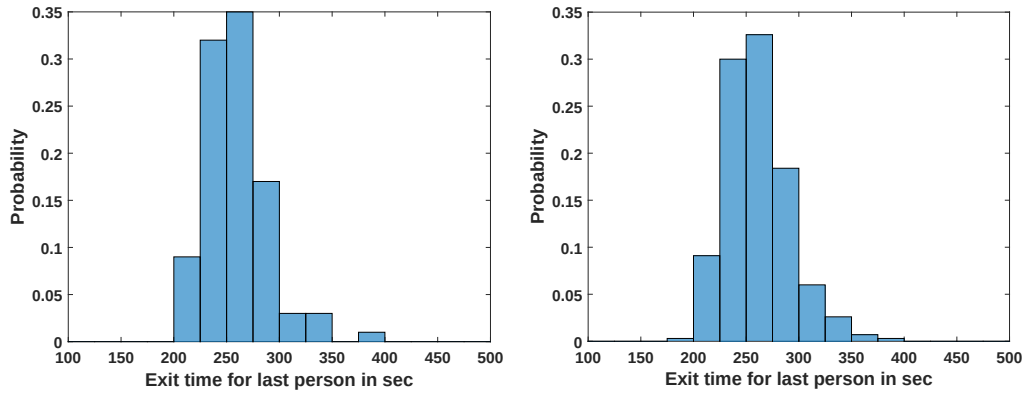
Figure 3.4: Illustration of interaction with spatially bounded confidence model

range (r) and modify their perceived value of states according to

$$V_{self} = (1 - \mu) \times V_{self} + \mu \times average\ of\ V_{others\ within\ r} \tag{3.2}$$

where μ is the herding level, which is how much weight individuals give to the herd's opinion. The value function is normalized for each individual to ensure that the herding effect is uniform.

An interaction process for an individual (blue) is depicted in Fig. 3.4. The boundary for the interaction/herding zone is shown with the green circle. Agents within the zone and not separated by walls are allowed to share opinion (green). With this combined movement, decision, and interaction model setup, various parameters and conditions affecting the egress dynamics were studied. The results are presented and discussed in the following section.



(a) (b)

Figure 3.5: (a) 100 runs, and (b) 1000 runs (Common parameters: $\alpha = 0.05$, $\mu = 0.6$, $r = 10\ ft$, $N = 200$, and $\tau = 4s$)

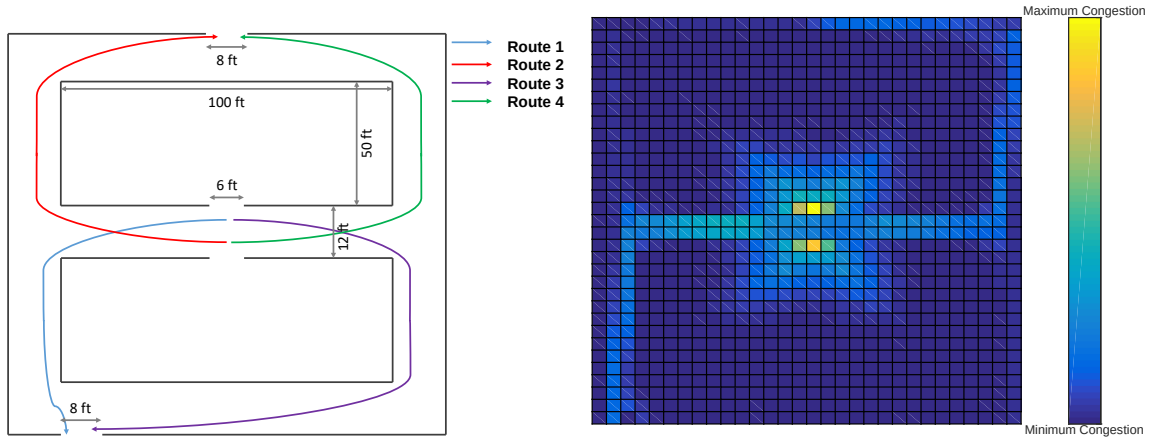
Table 3.1: Simulation parameters used for evaluating the minimum number of Monte Carlo simulation runs sufficient for extracting reliable statistics

Herding level	Herding range	Impatience growth level	Decision timer mean	Total number of people	Number of runs
$\mu = 0.6$	$r = 10ft$	$\alpha = 0.05$	$\tau = 4\ sec$	$N = 200$	100, 1000

3.3 Results and Discussion

3.3.1 Effect of number of runs

First, the effect of number of runs on the statistics was studied to find the minimum number of Monte Carlo simulations to get a reliable result. A set of Monte Carlo simulations with parameters given in Table 6.1 were conducted. A probability distribution for exit time of the last person from the building was obtained for each case (corresponding to 100 and 1000 runs) and shown in Fig. 3.5a and Fig. 3.5b respectively. The two-sample Kolmogorov-Smirnov test ([84]), a nonparametric hypothesis test was utilized to test whether both distribution came from the same cumulative distribution function at 1% significance level. The null hypothesis, both data came from the same distribution, was not rejected and hence the minimum number of runs was fixed at 100 for all the subsequent Monte Carlo simulations to extract reliable statistics. These tests were conducted with rational reward/reinforcement function. A detailed parametric study with rational decision makers is presented in the next section.



(a) (b)

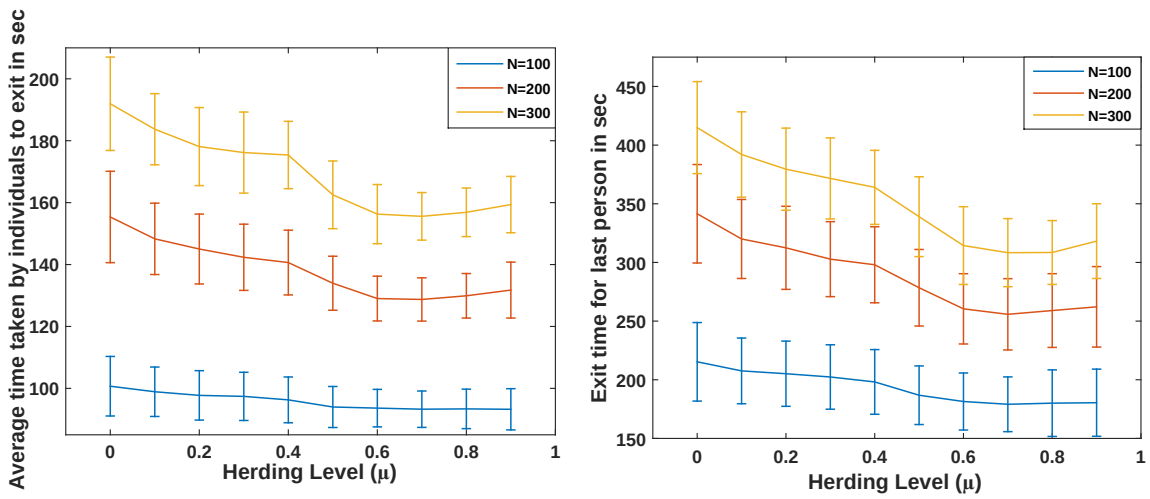
Figure 3.6: (a) Different exit routes available to the individuals, and (b) Heat map indicating congestion along the routes

3.3.2 Rational decision makers

For this set of simulations, the crowd's decision model was assigned a rational reward/reinforcement function. The reinforcement function was designed to reflect the path length for the various routes. The various routes available to the crowd are illustrated in the Fig. 3.6 along with the congestion map. As evident from the congestion map, route 1 (left, then down) was the most utilized path and route 4 (right, then up) was the second most utilized path. Route 1 was the natural choice of the rational informed crowd since it is the shortest path to safety. As every individual tried to go through route 1 it became crowded, impatience grew resulting in part of the crowd starting to move along route 4. The highest congestion occured at the room exits followed by the corridor just outside the rooms. The effects of herding behavior, frequency of decision making, and impatience level of the individuals under different total population size (N) are presented below.

Effect of herding level (μ)

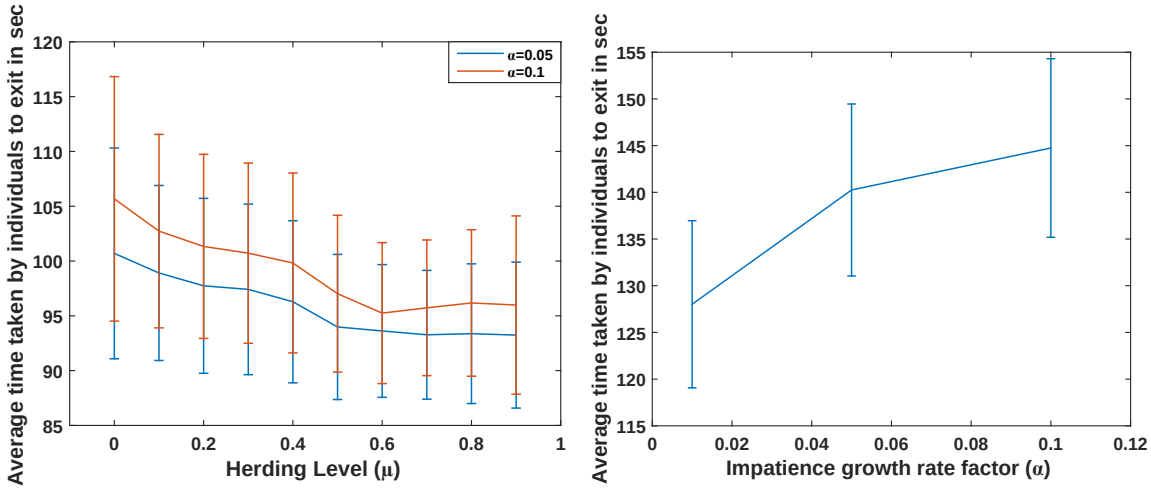
In this section the effect of listening to others' opinion is studied. Herding level (μ) determines the level of dependency on others' opinion. A herding level of $\mu = 0$ means the crowd doesn't depend on one another for their decisions. A herding level of $\mu = 0.4$ means every individual gives 40% weight to others' (within the confidence boundary) opinion and 60% weight to



(a) (b)

Figure 3.7: (a) Average time taken by individuals to exit the building with different herding levels (μ), and (b) Time when the last person has exited the building with different herding levels (μ) (Common parameters: $\alpha = 0.05$, $\tau = 4s$, and $r = 10\ ft$)

their own opinion. When an individual gives more weight to others' exit choice it correlates to stronger herding behavior. Fig. 3.7 depicts the average time taken by individuals to exit the building under different herding level and also the time when the last person exited the building. A herding range (r) of 10 ft with impatience growth rate (α) fixed at 0.05 was utilized for this set of simulations. Additionally, the decision timer distribution was set with a mean (τ) of 4 seconds and a standard deviation of 1 second; i.e., each agent revaluates their route choice every 4 seconds on average. When the total population (N) is lower, more herding led to quicker egress ($N = 100$), since information was shared and futile plans were quickly eliminated. Conversely, with a higher population ($N = 200$ and 300) more herding became detrimental with respect to the average exit time for individuals. The reason for this is higher herding level with increased population size led to elevated crowd density. It in turn contributed to a higher probability of becoming impatient and consequently resulted in increased route switching. From Fig. 3.7 , the trend of the average time to exit the building and time when the last person exited the building are qualitatively similar. Since the findings are qualitatively described on subsequent parametric studies, only the average time to exit the building is shown.

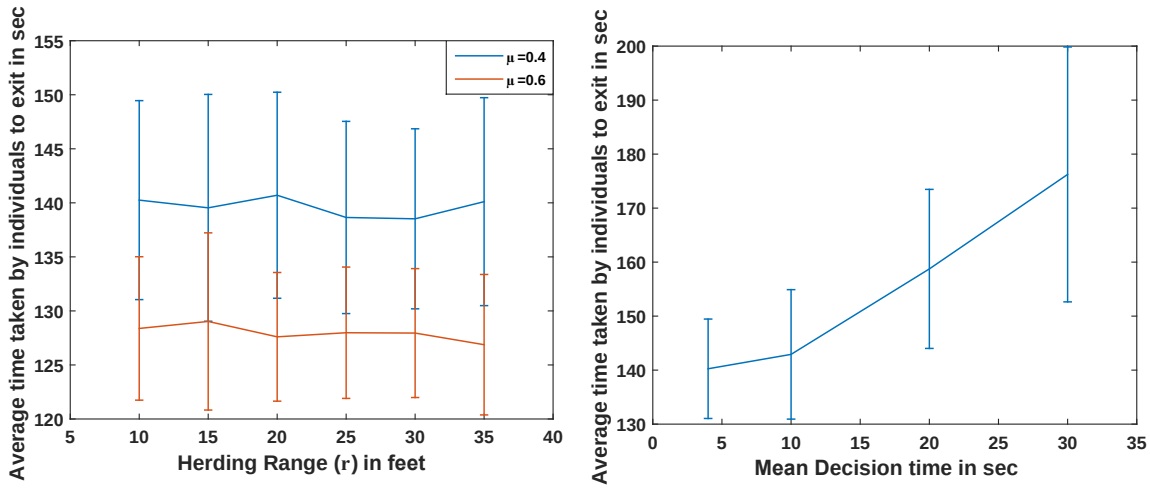


(a) (b)

Figure 3.8: (a) $N = 100$ - Combined effect of impatience growth rate (α) and herding level (μ), and (b) $N = 200$ and $\mu = 0.4$ - Effect of impatience growth rate (α) on the average time taken by individuals to exit the building (Common parameters: $\tau = 4\ s$ and $r = 10\ ft$)

Effect of impatience growth rate (α)

The effect of impatience growth rate (α) on the egress dynamics was investigated in this section. In a crowd with individuals possessing high level of impatience, a small bottleneck can lead to increased decision switching. To study the combined effect of impatience growth rate and different herding levels on an individual's average time to evacuate, the first set of simulations were conducted with population, $N = 100$, herding range, $r = 10\ ft$, and the decision timer was distributed as in the previous case. Figure 3.8(a) illustrates the effect of different impatience growth rate on same population. With a faster growth of impatience, the positive effect of herding on the evacuation time is lost at higher herding levels. This can be explained by higher probability of exit choice change by individuals when their impatience saturates faster. At a slower impatience growth rate, the crowd tend to stick with their initial exit choice for a longer time which implies lesser changes in exit choice of the crowd. Fewer changes in exit choice led to less time within the building. Thus, the crowd evacuated the building quicker. Isolating the effect of impatience growth rate in the second set of experiments with population, $N = 200$, herding level (μ) fixed at 0.4, and other parameters as in the first set faster impatience growth rate led to increased average time to evacuate the building (Fig. 3.8(b)).



(a) (b)

Figure 3.9: (a) $\tau = 4s$ - Effect of herding range (r) at two different herding level (μ), and (b) $\mu = 0.4$ and $r = 10\ ft$ - Effect of a different decision timer mean (τ) on the average time taken by individuals to exit the building (Common parameters: $\alpha = 0.05$ and $N = 200$)

Effect of herding range (r) and decision time

Next, the effect of herding range (r) and decision time of individuals (τ) on the evacuation time were studied in detail. Herding range (r) does not have any discernible effect on the evacuation time (Fig. 3.9(a)) with a rational reinforcement function. The individuals with a rational reinforcement function will have the same opinion towards the exits; hence, the herding range doesn't affect the evacuation time. These experiments were conducted with a crowd population, $N = 200$, an impatience growth rate, $\alpha = 0.05$, and a decision time mean (τ) fixed at $4s$. With the same set of parameters and with a herding level, $\mu = 0.4$, and a herding radius, $r = 10\ ft$, the effect of frequency of decision making on the crowd egress dynamics was examined. It is evident from Fig. 3.9(b) that the more frequent the crowd revaluates its decision the better it is for the evacuation time. A crowd of individuals who do not decide frequently can be stuck in a bottleneck or with an obsolete decision for a longer period of time. Additionally, even if an individual becomes impatient, the switching will not happen until the next decision cycle. Therefore, it is better for an individual to reconsider their previous exit choice frequently.

3.3.3 Biased decision makers

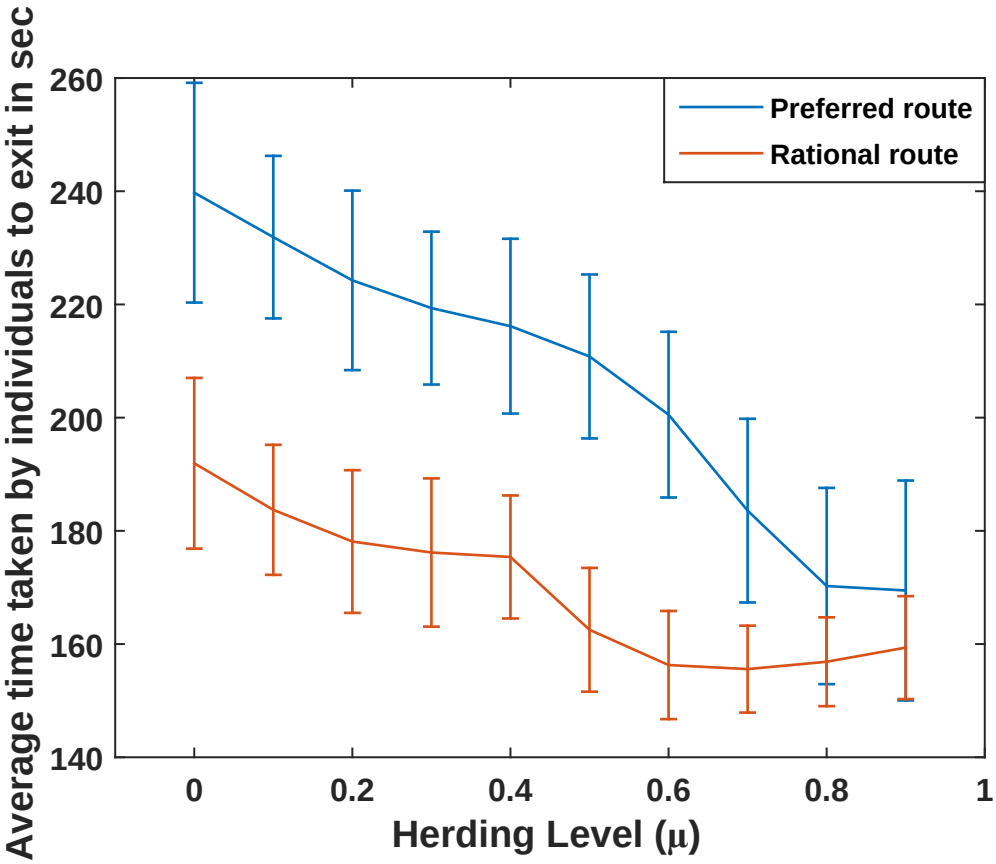
For the next set of simulation, the crowd was initialized with a biased reinforcement function. The crowd was evenly divided into four groups and each group was given a reward function that made one of the four available paths as the desired route for evacuation for the individuals in the group. People have a bias towards a familiar path and that is modeled through this biased reinforcement function. A crowd of 300 people were generated with an impatience growth rate, $\alpha = 0.05$. The crowd made decisions frequently (4 s) and the herding range was fixed at 10 ft. At all herding levels, the rational crowd fared better than the biased crowd (Fig. 3.10). Quicker evacuation was observed when the crowd consisted of more receptive individuals. Cooperation was better when individuals did not have complete unbiased knowledge of their environment.

3.3.4 Rational decision maker with biased leaders

The last set of simulations were conducted to study the effect of leaders with biased route choice on the crowd's egress dynamics. A leader is characterized by having a strong bias towards a particular route. Leaders are vocal and propagate their opinion in the crowd. To be seen as consistent, the leaders keep their opinion. The exit choice of a leader is affected only by the environment and not by other individuals. The crowd is composed of a few leaders and many rational decision making individuals.

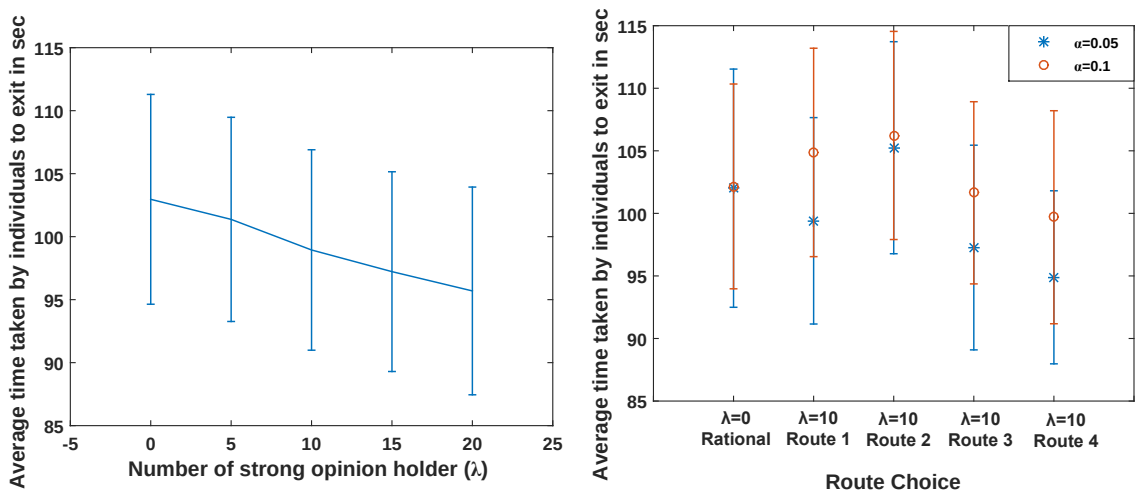
Effect of number of biased leaders (λ)

First set of simulations were conducted with λ number of leaders having strong inclination towards route 4 in a crowd of 120 people. The impatience growth level of the crowd was set at 0.05. The herding circle range was kept at 10 ft along with the herding level at 0.4 and the decision timer mean 4 s. The results are shown in Fig. 3.11(a). The average time to exit the building decreased with more leaders in the crowd. The crowd herded with the leaders and avoided congestion at route 1 and reached safety faster. Route 4 was chosen in particular because it was the second best choice among the available routes taking into account distance to travel and the potential bottleneck at exit 2.



(a)

Figure 3.10: Effect of different herding level (μ) on the average time taken by individuals to exit the building with rational and biased crowds (Common parameters: $N = 300$, $r = 10\ ft$, $\alpha = 0.05$, and $\tau = 4s$)



(a) (b)

Figure 3.11: (a) $\alpha = 0.05$, $N = 120$, and leader with route choice 4 - Effect of number of strong opinion holders on the average time to evacuate, and (b) Number of strong opinion holders, $\lambda = 10$ - Effect of different route choice of leaders on the average time taken by individuals to exit the building (Common parameters: $\tau = 4s$, $r = 10\ ft$, and $\mu = 0.4$)

Effect of route choice of biased leaders under different impatience levels

The final set of simulations were concerned about the route choice of the leaders. The simulations were conducted with fixed number of leaders ($\lambda = 10$) in a crowd of 110 people. The herding level, herding range, and decision timer were the same as in the previous case. The effect of leaders were diminished (Fig. 3.11(b)) when a crowd consisted of individuals with faster impatience growth ($\alpha = 0.1$). The switching of exits was more prevalent in a highly impatient crowd, leaders' influence was less impactful, which led to an increased egress time. With a lesser impatient crowd ($\alpha = 0.05$), except for route 2 which puts additional pressure on already crowded lane all other leaders route bias were helpful in getting the crowd to safety quicker. Even leaders with bias towards the shortest route had a positive effect. The leaders were able to keep the crowd directed towards the shortest route for longer time even if they became impatient due to bottlenecks and crowded lanes.

3.4 Summary

This part of the research combines a naturalistic movement model and a decision making model with opinion sharing dynamics (the hybrid model) to study the effect of opinion sharing on the crowd evacuation metrics. We found that the effect of opinion sharing is dependent on the state of the crowd. Factors such as how receptive the crowd is to opinion sharing, how fast the individuals tend to change their exit choice when confronted with crowded lanes/bottlenecks, and the frequency of decision making affect the crowd's evacuation time. Ideally, a tolerant rational crowd with well informed leaders/strong opinion holders is well-suited for a quick evacuation of a building. Herding is not detrimental for evacuation. However, over-herding can lead to under utilization of all the available routes and an increase in the evacuation time. People with a strong opinion can help with faster egress if their strong opinion aligns with the under-utilized route(s). If the overall state of the crowd is calm, which lends itself to a better propagation of opinion, then it helps the crowd to exit the building quicker. We have presented a simulation model that combines opinion sharing with a movement and a decision model in this chapter. In future work we intended to collect experimental data to corroborate our simulation results.

Chapter 4

Reward learning with Inverse Reinforcement Learning algorithm

4.1 Objective

In the previous chapter (Ch. 3), Markov decision process is the underlying model for determining the instantaneous exit choice. The reward function defined by the expert played a crucial role in the route chosen by the simulated individuals. In other words, reward function is the most succinct representation of the underlying decision mechanism [85]. From the decision theoretic perspective, in the forward problem, an expert specifies the reward function and the optimal value of state-action pair(expected cumulative reward) are determined which leads to an optimal policy (a state-action map). This was the case in the research carried out in the preceding chapter. Conversely, the inverse problem involves an expert demonstrating a policy and the agent recovering the hidden reward function to explain the expert's behavior as the optimal policy. This falls under learning from demonstration paradigm and is formally called inverse reinforcement learning [85].

In the preceding chapter, a naive reward function based on shortest exit route was utilized. The problem was formulated as a forward learning process. From the evacuation model point of view, it is desirable to extract the reward function from a demonstration since it is easier for a person to demonstrate an escape plan rather than explicitly specify the internal reward

function that takes into account several factors that played a role in their decision process*. This is a typical inverse learning problem and a brief history of several techniques popular in the field is presented below.

4.2 Literature review

Learning from demonstration is an interesting paradigm usually studied from the robotics context and has been tackled by many researchers. Most of the prior work have tried to address it from the viewpoint of database building and searching in the database for the current situation and executing the script from the database [86]. Initiated in late 1980s as imitation learning, the target of early research in reinforcement learning was to make manipulators follow similar path from start to goal as previously demonstrated by an expert. Segre and Dejong [87] extracted a set of 'if-then' sequences to achieve the path imitation. Given the limitations of available computing resources in the late $80's$, this itself was a compelling feat. As the computing power and sensor technology continued to improve, researchers began to develop systems that are more intelligent. Latest imitation learning technique as reported in [88] tries to incorporate both position and force profile into the learning domain. Another work [89] tries to use Gaussian Mixture Model and Gaussian Mixture Regression to learn the way-points to either lead/follow in the task of picking up an object alongside a human. Another group has trained a manipulator both in simulation and in real-time to catch a flying object[90, 91].

Another body of work by Veloso's group introduced a new method called confidence based autonomy[92, 93]. The basic building block of their algorithm was a robust database where each distinct state action pair is stored. In real-time execution when a state is encountered by the agent, it queries the database for a suitable action. The database returns a recommended action along with a confidence parameter. If the confidence is below a set threshold then the agent request for a demonstration.

*Results presented in this section are from our published work: Srinivasan, A. R., & Chakraborty, S. (2016, August). Path planning with user route preference-A reward surface approximation approach using orthogonal Legendre polynomials. In Automation Science and Engineering (CASE), 2016 IEEE International Conference on (pp. 1100-1105). IEEE.

Each of the techniques for learning from demonstration described above has its own unique advantages and disadvantages. The problem with database-oriented technique is the storage of all the relevant information from training in an intelligent manner for it to be quickly accessible. If the information becomes too large then real-time fetching will become time consuming. There is a similar state space explosion problem associated with Markov decision processes. The time to find the optimal solution scales exponentially with the number of distinct states in Markovian world.

There has been a body of work by Ng's group [94, 95, 85] on modeling the learning problem as a Markovian process. The demonstrations are assumed to be manifestation of the expert's policy, which is considered as the optimal solution to the implicit Markov Decision Process (MDP) with unknown reward functions. The inverse reinforcement learning algorithm is used to compute the unknown reward function from the expert's demonstration(s). In the work by Kim et.al. [96, 97], the path planning with human input is accomplished by hand-picking a set of features and learning the weights for each feature by using inverse reinforcement learning. Similarly [98] attempts to incorporate human factor into autonomous path planning by selecting specific features from the sensor input. The pros and cons of different feature sets are dealt with in [99]. The failed set of demonstration were used in [100]. Nguyen et al. [101] splits the state space into different region and computed the augmented reward function by utilizing expectation maximization technique. Ziebart et al. [102] utilized maximum entropy method to learn and predict user's route preference and destination. There is also a work by Deisenroth et al. [103] wherein they try to account for incomplete models. In all of these works, domain expertise is required in order to hand pick the feature set.

In this work, we are also trying to model the agent as an MDP with unknown reward functions to be learned from demonstration(s). The difference from the previous work is that we are trying to circumvent the need for domain knowledge and hand picking the feature set by utilizing the orthogonal polynomial functions as basis functions (the feature set) for representing the reward structure. Additionally, we can circumvent the problem of state space explosion by utilizing polynomial function of order lower than that of the state space.

This is largely inspired by image reconstruction techniques employed in image processing community [104].

A model experiment consisting of a tele-operated robot in an arena was designed to test the modified inverse reinforcement learning algorithm. The mathematical background and the algorithm are presented in the following section.

4.3 Mathematical Background

This work employs two underlying principles, namely Markov decision process [72, 105] and inverse reinforcement learning [94, 95, 85] developed by Ng's group. A succinct description of both is provided here for clarity and completeness.

4.3.1 Markov Decision Process

A Markov decision process $M = \{S, A, P, \gamma, R\}$ consists of the following

S Set of all possible states of the system.

A Set of actions available to the system.

P Transition probability $\mathrm{P}(s, a, s')$ which gives the probability of transition to state s' from state s by taking action a.

R Set of rewards - This indicates the payoff from the various states of the system. The system's overall behavior depends on the rewards.

γ Discount factor $\in [0, 1)$ - This parameter controls the relative weights of rewards acquired in near vs. distant future.

The basic underlying assumption of a Markov decision process (MDP) is that the current state and the action taken alone determines the next state, independent of past states or actions. For a MDP, the policy, π is a prescription of action(s) to be taken from given states. A policy is optimal, if it satisfies the Bellman optimality equation. To describe the optimality equation, value function V and Q function have to be defined. Let a MDP

$M = \{S, A, P_(sa), \gamma, R\}$ and a policy $\pi : S \rightarrow$ A be given. Then, for all $s \in S$, $a \in A$, the value function V^π and Q function (Q^π) have to satisfy

$$V^\pi(s) = R(s) + \gamma \Sigma_{s'} \left(P(s, a, s')\right) V^\pi(s') \tag{4.1}$$

$$Q^\pi(s, a) = R(s) + \gamma \Sigma_{s'} \left(P(s, a, s')\right) Q^\pi(s', \pi(s')) \tag{4.2}$$

The value function and Q function represent the expected cumulative reward for following the given policy π and a policy π is an optimal policy π^* for M if and only if $\forall s \in S$,

$$\pi(s) \in arg\ max_{a \in A} Q^\pi(s, a) \tag{4.3}$$

This simply states that at any given state, the action chosen must result in the system being in the best possible next state with respect to their calculated value.

4.3.2 Inverse Reinforcement Learning

A typical well-defined Markov decision process problem consists of a set of all possible states (S) and action (A), the transition probabilities of each state-action pair (P), the discount factor (γ) and a reward function (R). Given this 5-tuple, the aim of a Markov decision algorithm is to find a policy that maximizes the total reward obtained from the start state(s) to the goal state(s). A policy that maximizes the total collected reward is called an optimal policy. The linchpin of the entire process is the specification of the reward function.

In the inverse problem, the agent does not have direct access to the underlying reward function, but is only shown positive examples of how a task might be performed. The assumption is that the demonstrator has an implicit reward function and the demonstration is a manifestation of the optimal policy with respect to that reward function. The inverse reinforcement learning problem deals with extracting the reward function that best explains the policy demonstrated by the expert.

We restrict ourselves to the case of $S = \mathbb{R}^2$, for example, longitude and latitude can completely specify intersections. If we consider the state space to be 2-dimensional then the

reward function computed by the inverse reinforcement learning algorithm has to map from $\mathbb{R}^2 \longrightarrow \mathbb{R}$. Considering the difficulty of optimizing over this space, a linear approximation for the reward function can be used, where

$$R(s) = \alpha_1\phi_1(s) + \alpha_2\phi_2(s) + \alpha_3\phi_3(s) + \ldots \alpha_n\phi_n(s) \tag{4.4}$$

In [94][85], for the linear approximation of the reward function, R the authors had hand picked the feature set. The same is the case in existing techniques for user to input their route preference[97, 96, 98, 101]. However, if no such insight is available, a simple but impractical set of basis functions with the same dimensionality as the number of states can be constructed as follows. For instance, an example basis function array for a space discretized into $2 \times 2 = 4$ distinct states can be

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \text{and} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{4.5}$$

where each matrix represent one of the basis function. This is the simplest of basis function array which can represent any reward shape in 2D for the 2×2 state space. But it is evident that with increasing number of states this will lead to exponential increase in computation time for the inverse algorithm.

To alleviate the problem, we take inspiration from the image processing community [104], where multivariate orthogonal polynomials are used as basis functions to find the image moments. One discrete orthogonal polynomial function that has been tested with success is Legendre polynomial of different orders. A Legendre polynomial is given by

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} \left[(x^2 - 1)^n\right] \tag{4.6}$$

where n denotes the order of the polynomial.

The reward function can be considered as a complex envelope encompassing the entire state space. To find the equations governing that envelope, utilizing a set of orthogonal polynomials reduces the number of variables to be optimized. The only variables that need

to be optimized are a fixed number of coefficients of the orthogonal polynomials, regardless of the size of the state space. The orthogonality of the polynomial function allows us to compute the coefficients for each dimensions separately and then use tensor product to find the value for a given (x,y) coordinate. This is evident from the reward envelope (shown in Fig. 4.1(b)) found by the modified algorithm for the tele-operated robot navigation. The smooth surface of the reward function is the result of using the weighted sum of orthogonal polynomial basis function to approximate the original implicit reward function. If we approximate the reward function, R with Legendre polynomials, then R is given by

$$R(s) = \alpha_1\phi_1\theta_1 + \alpha_2\phi_1\theta_2 + \alpha_3\phi_1\theta_3 + \ldots \alpha_{n\times n}\phi_n\theta_n \tag{4.7}$$

where n is order of Legendre polynomial and θ and ϕ are the Legendre polynomials of various orders, one for each dimension. The α_is are the parameter our inverse reinforcement learning algorithm is attempting to optimize. Since expectation is a linear function, the value function, V corresponding to the reward function, R given by equation (4.7) is

$$V^\pi = \alpha_1 V_1^\pi + \alpha_2 V_2^\pi + \cdots + \alpha_{n\times n} V_{n\times n}^\pi \tag{4.8}$$

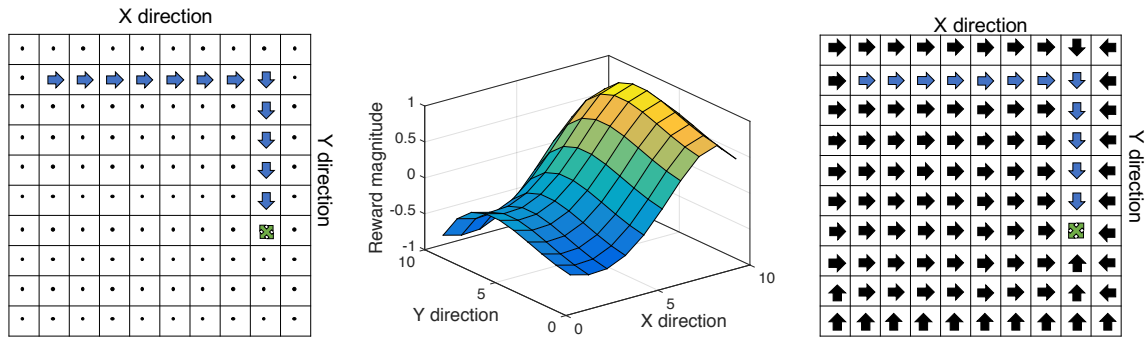
Thus Bellman's optimality equation (4.3) can be written as

$$E_{s'\sim P_{sa_1}}[V^\pi(s')] \geq E_{s'\sim P_{sa}}[V^\pi(s')] \tag{4.9}$$

for all states s and all actions $a \in A \setminus a_1$. This merely states the Bellman equation (4.3) in another form. From equation (4.8), we know that $V^\pi(s)$ is a linear combination of basis function weighted by α_is. Hence we can formulate the problem as linear programming (LP) to find the constraints (α_is). We utilize the linear programming formulation from Ng and Russell's work [85]

$$\text{maximize} \sum_{j=1}^{k} p\left(\hat{V}^{\pi^*}(s_0) - \hat{V}^{\pi_j}(s_0)\right) \tag{4.10}$$

s.t.$|\alpha_i| \leq 1, i = 1, \ldots, n \times n$ and j is number of iteration algorithm has gone through so far



(a) (b) (c)

Figure 4.1: (a) The path demonstrated to the Turtlebot, (b) The extracted reward for the path and, (c) The optimal policy extracted from the reward function

The α_i comes into play through (4.8) and the penalty function used here is given by $p(x) = x$ if $x \geq 0$, $p(x) = 2x$ otherwise.

The current algorithm as presented in [106] modified from [94, 85] to suit the tele-operated system(s) is elucidated in Algorithm 4.

Algorithm 4: Modified inverse reinforcement learning algorithm

Data: $S, A, P, \pi_{Expert}, \gamma$
Result: R

```
Initialize with a set of basis functions. A set of Legendre polynomial with fixed order is
  chosen in this work.
Calculate the value of the states using value iteration algorithm for the expert's policy.
Randomly pick a policy and add it to set of policies. (A random policy is used to seed
  the algorithm)
while Reward function satisfying the expert's policy is not obtained do
    Calculate the value of the states using value iteration algorithm with each of the
      basis function for all the policies in the set.
    Maximize the weighted difference between the expert's policy value and the
      average value from the set of policies.
    Use the coefficients to find the new reward function.
    Compute the Q-Value, find the respective policy for the reward function, and add
      it to the set containing the random initial policy.
end
```

The weighted difference between expert's policy and average value from the all other policies in the set is maximized, in a sense we are trying to find a reward function that maximally differentiates between expert's policy and all other possible policies. The

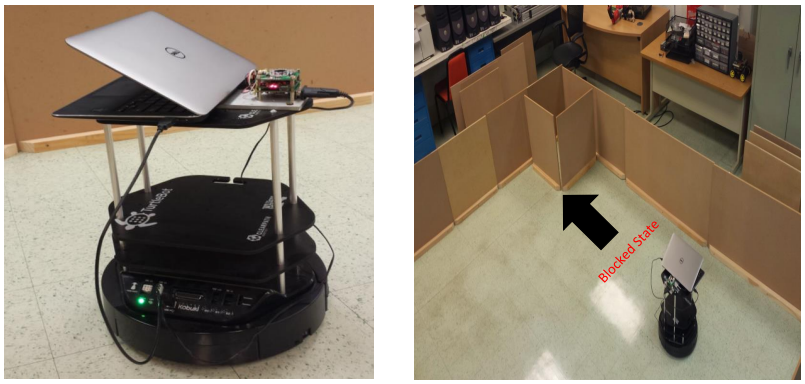
extracted weight/reward function can be utilized to find the complete policy of the expert. The order of the polynomial is found by starting with order 2 and increasing in steps of 1 till a sufficient representation of reward function is achieved. In our test case with 100 distinct states in 2D space, a pair of Legendre polynomial with order 6 was sufficient to find reward function for all of the test paths. Thus instead of a maximization problem posed over 100 coefficients, it is reduced to only 36 (6×6) coefficients. Thus we circumvent the state space explosion problem by utilizing orthogonal polynomials of an order much lower in comparison to the number of distinct states in the system.

4.4 Experimental Setup

The experimental setup for the path-planning robot consist of a Turtlebot and a stargazer indoor GPS system. The Turtlebot is a low cost robot kit which runs on open source software ROS (Robot Operating System). The stargazer is a low cost indoor GPS which works on the principle of infrared image processing. Markers on the ceiling are read by an infrared camera on the stargazer and analyzed on board to provide the estimates of current position and orientation of the Turtlebot.

A point to be noted is the data from the stargazer is prone to noise. The same has manifested itself as random points in the reconstructed path. Also the stargazer sensor has been mounted off-center on the Turtlebot (figure 4.2(a)) which has lead to small loops in the reconstructed trajectories wherever the Turtlebot was making turns. The work flow can be simply stated as follows. First, a demonstration from an expert is recorded. The state space is divided into rectangular grids and from the recorded demonstration the state-actions pair are interpreted. Then the modified inverse reinforcement learning algorithm is run on the available data and once a suitable expert policy is extracted, the algorithm is stopped and the policy is fed back to the autonomous agent.

- The first experiment was designed to show that the Turtlebot can acquire the human demonstrated path and follow the same in the autonomous mode. The state space has been defined as twnty five equal sized squares on the arena floor. The action for the Turtlebot are restricted to rotate left, rotate right, move forward, move backward



(a) (b)

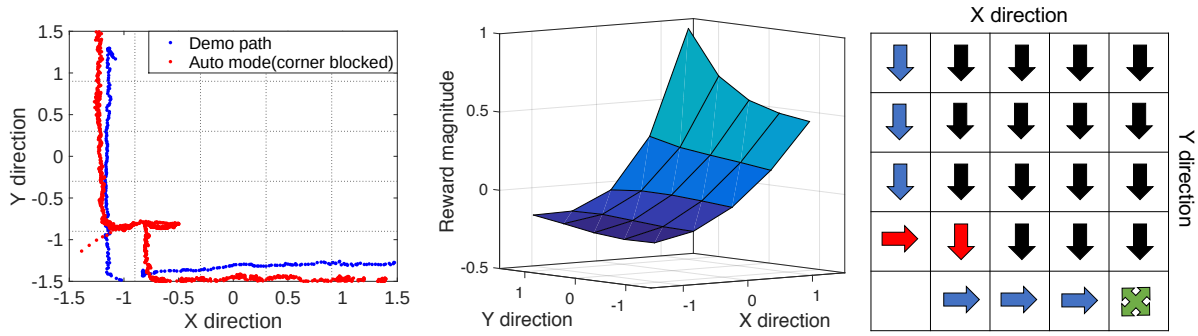
Figure 4.2: (a) The Turtlebot platform equipped with a Stargazer indoor GPS. (b) Turtlebot in the arena. A corner in the arena is blocked to test the ability to adaptively re-plan.

and halt. Once a demonstration is recorded, the GPS data are utilized to extract the states and the state transition in the demonstrated path. Then the modified inverse reinforcement learning algorithm is run and the expert's unknown reward function and the complete policy is extracted.

  - As a next step, a corner that comes in the path is cordoned off and the ability of the algorithm to come up with an alternate policy which matches the expert's path as much as physically possible is tested. For this step, the state transition into the blocked corner is voided.

- The next experiment is to demonstrate a complex path to the Turtlebot and then once a policy is extracted by the algorithm, the Turtlebot is started from a different start point to test the ability of the robot to still follow the expert's demonstrated path. This experiment was designed to showcase the ability of the algorithm to extract a reward function for a complex policy and also reach the destination from a different start point and match the expert's policy in an intelligent way.

- The last set of experiments is done to show the advantage of utilizing the polynomial basis function. For this, the complex path (path with maximum number of permissible turns) is taken. The learning algorithm is run for different number of distinct states with both the simple basis function set (has dimensionality equal to the number of



(a) (b) (c)

Figure 4.3: (a) The demonstrated path (blue) and the path followed by the Turtlebot in autonomous mode (red) when a corner in the demonstrated path is made inaccessible, (b) The extracted reward for the demonstrated path and, (c) The optimal policy extracted from the reward function

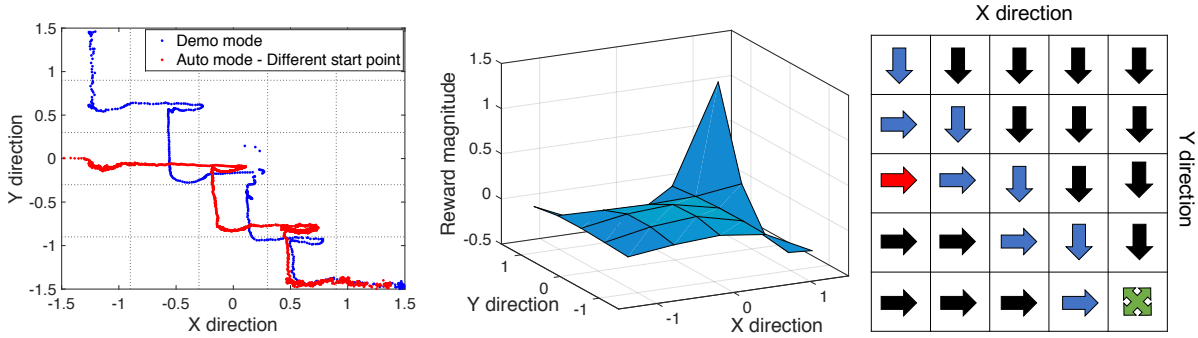
states) and polynomial basis function (fixed number of coefficients regardless of the number of states). The time complexity graphs showing the results are generated.

4.5 Results and discussion

Figure 4.3 shows the path demonstrated by an expert to the Turtlebot (in blue). The path followed in autonomous mode after the policy is extracted using the inverse reinforcement learning algorithm is similar to the demonstrated path, thus validating that the extracted policy tries to mirror expert's path. The arrow in the policy graph corresponds to the desired direction of movement as extracted by the algorithm. Figure 4.3 shows the ability of the robot to maneuver the cordoned off corner and follow the expert's path as much as physically possible (shown in red).

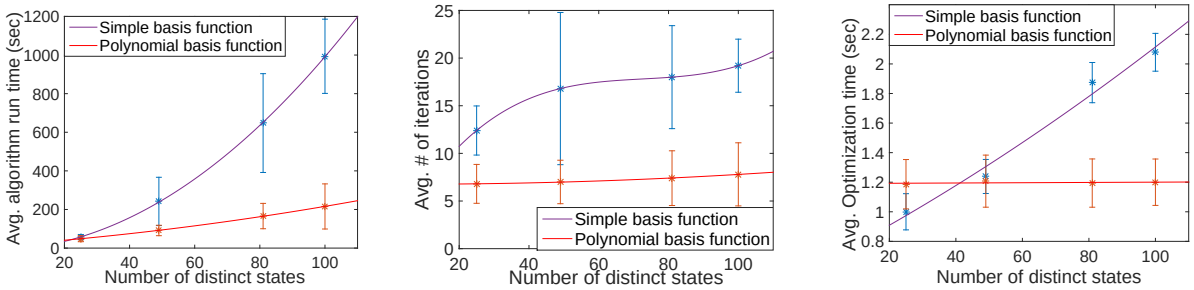
Figure 4.4 shows the ability of the robot to follow even a complex path from a different starting point. It may be noted that the learned policy tries to keep to as much of the demonstrated path as possible. In other words, even from a different starting point, the robot joins the demonstrated path as quickly as it can without violating any physical constraints.

Figure 4.5 shows required computation times for different number of distinct states. Figure 4.5(a) shows that the average time to run the complete learning algorithm with the simple basis function increases exponential with the number of distinct state. Whereas the



(a) (b) (c)

Figure 4.4: (a) The path demonstrated and followed from a different starting point by the Turtlebot, (b) The extracted reward for the path and, (c) The optimal policy extracted from the reward function



(a) (b) (c)

Figure 4.5: (a) Number of distinct states vs. average time taken for the learning algorithm to find the expert's reward function, (b) Average number of iterations taken by the algorithm to find the expert's reward function, (c) The average time taken for the optimizer to find a solution

average run time with the polynomial basis function is almost linear with the number of distinct state. This is result of constant number of variables to be optimized in case of the polynomial basis function compared to increasing number of optimization variables in case of the simple basis function.

The linear increase in polynomial basis function case is the result of running value iteration for increased number of states. Figure 4.5(b) shows the average number of iterations required for the algorithm to find the expert's implicit reward function. Figure 4.5(c) depicts the time taken by just the optimization routine to find the solution for given set. Since the number of optimization variables is constant in the polynomial basis function case, the

optimization routine time does not change with increasing number of distinct states. But, in the case of simple basis function, the optimization routine time increases exponentially with number of distinct states and thus results in more running time for the entire learning algorithm. Time is a crucial factor when running real time systems and the graphs prove that it is advantageous to approximate the reward function using polynomial basis functions.

4.6 Summary

Thus the expert/user can provide a demonstration to the agent, which is more natural than specifying the user preferences. From that demonstration the underlying implicit reward function for the user preferences can be extracted in a timely manner and utilized to understand the expert's behavior. The mental load on the expert to explicitly specify the reward function over the entire state space is removed. In future research this reward learning algorithm can be utilized to extract insights about pedestrian's route choice from a reward function perspective. This can enhance the closeness of the simulated decision model with actual human decision making process.

Chapter 5

Realistic estimation of building evacuation time

5.1 Objective

Emergency evacuation is a stressful situation. Generally, it is good to know an estimate of how long it will take for a building at its maximum occupancy to be evacuated. This minimum evacuation time estimate can help building planners to take into account appropriate design of the exit structures to avoid unrealistic minimum evacuation time. Starting with Evacnet+ [107], there are several optimal flow calculating algorithms available either as an academic project or commercial product to estimate the minimum time to evacuate a given building. Lin et al. [108] provide a more recent optimization algorithm for evacuation planning. Yusoff et al. [109] provide a comprehensive overview of existing techniques for evacuation optimization. Lu et al. [110] provide a couple of heuristic methods which give comparable results to Evacnet, which is standard tool to compute minimum time to evacuate a given building. Evacnet has an exponential run time which scale with the network and hence not a suitable candidate for large building evacuation simulation. In this work, we have utilized one of the heuristic algorithm elucidated by Lu et al. [110] as the baseline optimal evacuation strategy and a plan for a responsible individual (i.e.,) guide is extrapolated based on it.

5.2 Algorithm description

5.2.1 Network setup for testing the evacuation algorithm

A given structure as in figure 5.1 is converted into equivalent network consisting of nodes and edges as depicted in figure 5.2. All rooms, corridors, staircase and building exits are converted into equivalent nodes with respective maximum and initial occupancy specified. All the capacity constraint and travel time are depicted utilizing edges. With the converted building specified as nodes and edges with appropriate properties of the building, the heuristic optimal evacuation algorithm is first run on the graph. The heuristic optimal strategy is adopted from Lu et al. [110] and is reproduced below for completeness (Algorithm 5).

In the multiple route capacity constrained routing approach, one computes the next best available route at every instance and reserve the best available path (path with the shortest time to exit from any of the unevacuated source nodes) at every time instance. When a given path is reserved the capacity of the nodes and edges along the route is changed accordingly and thus when the next best available route is computed, the capacity constraint from the previous reservations is taken into account. When at a given time instance there is no more available path to safety, the time is incremented by 1 second and the process is continued till every evacuee has reserved a path to safely exit the building. This algorithm has been proven by Lu et al. [110] to produce comparable result to the benchmark, Evacnet and is scalable with increasing network size.

Next, each individual is assigned a preferred route by randomly picking one of the available route from their source node to exit the building at the start of the simulation. The evacuation time for nominal strategy is computed by randomly picking an individual starting from $t = 0$ and assigning/reserving their preferred route if it is available. When all the routes at a particular time, t is reserved, the time is incremented by 1 second. The reservation of nominal route is done till every person has reserved their preferred route. Next, the difference between the last person time to exit a given source node according to the optimal plan and the nominal plan is computed for every node. After that the realistic evacuation time is estimated by utilizing the following algorithm 6.

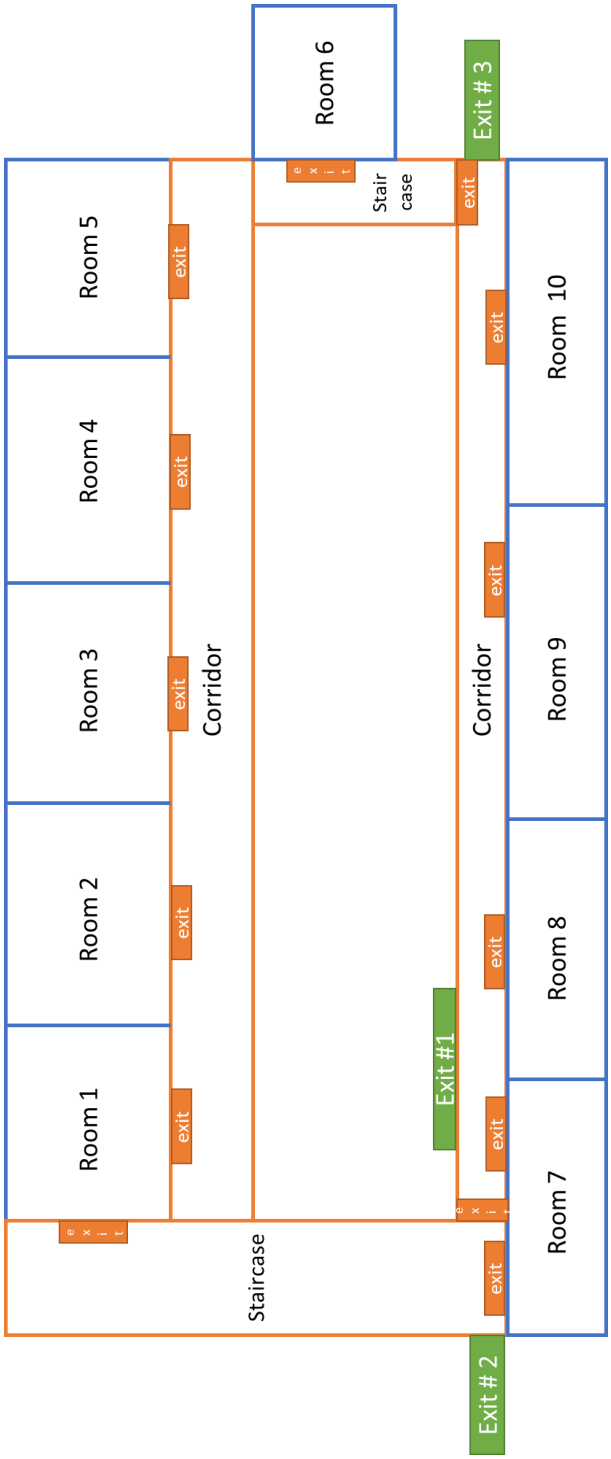


Figure 5.1: Illustration of the node-edge equivalent of sample building

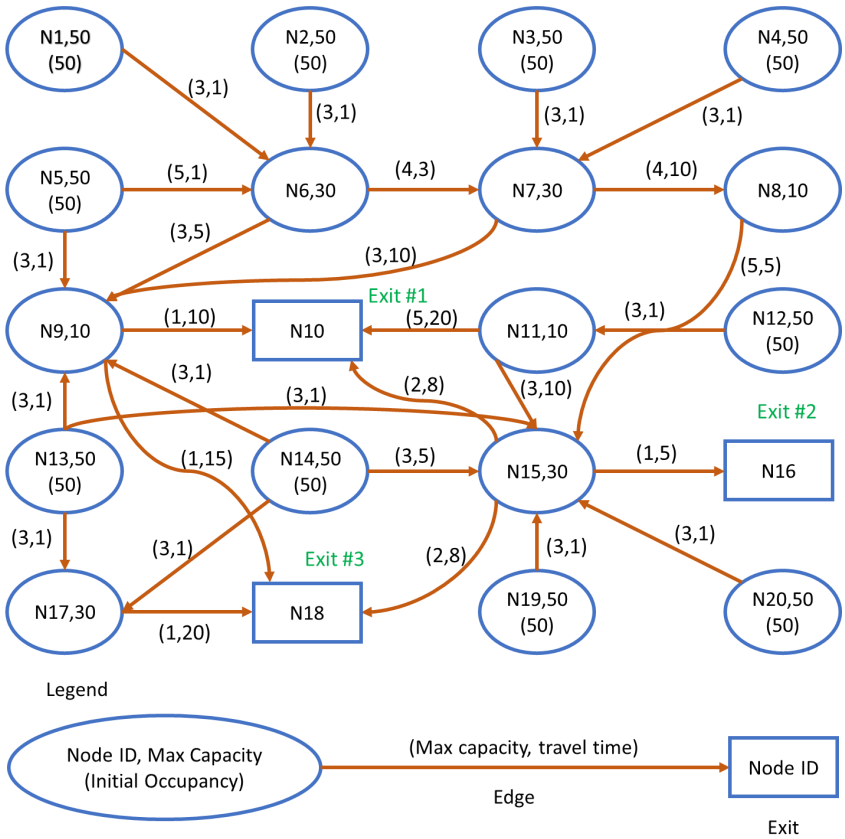


Figure 5.2: Illustration of the node-edge equivalent of sample building

Algorithm 5: Optimal evacuation strategy - Multiple Route Capacity Constrained Routing Approach (MRCCP) from Lu et al. [110]

Data:

1) $G(N, E)$: a graph G with a set of nodes N and a set of edges E;
   Each node $n \in N$ has two properties:
      Maximum node capacity(n) : non-negative integer
      Initial node occupancy(n) : non-negative integer
   Each edge $e \in E$ has two properties:
      Maximum edge capacity(e) : non-negative integer
      Travel time(e) : non-negative integer
2) S : set of source nodes, $S \subseteq N$;
3) D : set of destination nodes, $D \subseteq N$;

Result:

A heuristic optimal evacuation plan

```
while any source node s ∈ S has evacuee do
    find route R < n_0, n_1, ...., n_k >= with earliest destination arrival time among
     routes between all s, d pairs, where s ∈ S, d ∈ D, n_0 = s, n_k = d;
    flow = min( number of evacuee still at source node s,
                Available edge capacity (all edges on route R),
                Available node capacity (all nodes from n_1 to n_k on route R),
                    );
    for i = 0 to k − 1 do
    {
        t' = t+ Travel time (e_{n_i n_{i+1}});
        Available egde capacity (e_{n_i n_{i+1}}) reduced by flow;
        Avaiable node capacity (n_{i+1}, t') reduced by flow;
        t = t';
    }
end
Post-process results and save heuristic optimal evacuation plan;
```

Algorithm 6: Algorithm for finding the realistic evacuation time of a given building structure

Data:

1) Optimal evacuation strategy from Multiple-Route Capcity Constrained Routing Approach(MRCCP)

2) Nominal evacuation strategy using individuals' preferred/familiar/pre-determined path to exit the building

3) A start node for the responsible individual/guide

4) Difference in time between the last person to start from a given node according to optimal plan and the nominal plan

Result:

A realistic estimate of evacuation time and the path for responsible individual/guide to help crowd evacuate as close to the optimal plan as possible

```
while For all individuals who has not reserved a path to exit the building do
    if Guide has not visited the individual's starting position/node then
        Try to reserve the nominal path
    end
    if Guide has visited the individual's starting position/node then
        Reserve the optimal path for the individual (if path not available at current
          time, increase the time in increments of 1 second till the optimal path can be
          reserved)
    end
    Next node to visit is determined by comparing the time difference data and the
      worst offending node is selected. Guide is moved according to time constraint
      (from the network specification) to the selected node.
end
The realistic evacuation time is estimated from the guided path. (Both the guide's
  path and the guided path of individuals is saved)
```

Table 5.1: A sample of the heuristic optimal evacuation plan saved after running the MRCCP algorithm

Group of People					
ID	Origin	No. of People	Start Time	Route	Exit Time
A	N19	1	0	N19-N15-N16	6
B	N19	2	0	N19-N15-N10	9
C	N20	2	0	N20-N15-N18	9
D	N12	3	1	N12-N11-N10	22

The algorithm 6 presented above gives the current basic approach adapted to compute a realistic estimate of the minimum time to evacuate an occupied building. It is assumed that a guide/responsible person who can be a security personnel, building evacuation manager or an appropriately equipped person is moving through the building once emergency evacuation is necessitated. The guide is moving according to a specific plan which attempts to mimic the optimal evacuation strategy by visiting node/rooms in decreasing order of time difference between the optimal strategy and the nominal strategy of the source nodes. The underlying principle is to the stem the non-optimal flow starting from the worst offending source node and end with the least offending source.

5.3 Results and Discussion

The figure 5.3 shows the comparison chart between the heuristic optimal plan, the realistic evacuation plan and the average time if individuals just egress according to their preferred route choice. It is clear from the figure that with a guide moving in pseudo-optimized route starting from his/her initial node, they are able to provide valuable instruction to the evacuating crowd to produce comparable result to the optimal evacuation strategy. Table 5.1 present a sample of the optimal evacuation plan computed by the multi-route capacity constrained planner (MRCCP). It can be seen that MRCCP is trying to reserve the best path available for as many individuals to evacuate as possible in any given time frame. Table 5.2 shows the reservation made if everybody is trying to reserve their preferred path to safety. Finally, table 5.3 shows a portion of the guided path taken by the building occupants to safety taking into consideration the path and time constraint of the guide.

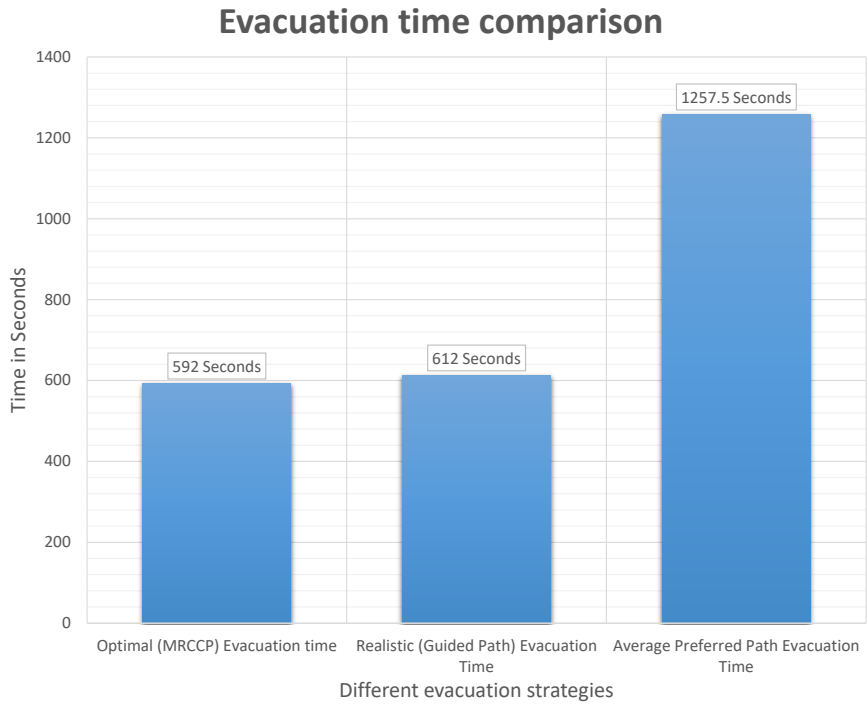


Figure 5.3: Comparison between different evacuation strategies

Table 5.2: A sample of the nominal evacuation plan computed by reserving individuals preferred path

Group of People			Start Time	Route	Exit Time
ID	Origin	No. of People			
A	N3	1	0	N3-N7-N9-N10	21
B	N1	1	0	N1-N6-N7-N8-N15-N18	27
C	N12	1	1	N12-N11-N10	22
D	N1	1	7	N1-N6-N7-N8-N15-N10	34

Table 5.3: A sample of the guided evacuation plan calculated taking into account the guide movement

Group of People			Start Time	Route	Exit Time
ID	Origin	No. of People			
A	N1	1	0	N1-N6-N7-N8-N15-N16	25
B	N1	2	0	N1-N6-N7-N8-N15-N10	28
C	N2	1	0	N2-N6-N7-N8-N15-N18	28
D	N13	1	1	N13-N15-N16	12

This realistic evacuation time estimator can be easily modified and morphed into a mobile application which can distributed in a real-life evacuation scenario. Utilizing this mobile application, a responsible person can help the crowd to deviate from their preferred path and closely follow the optimal evacuation strategy for the given building. Thus, it can lead to a visible and quantifiable improvement of the overall evacuation process.

5.4 Summary

In this part of the research, an existing optimal evacuation time estimator algorithm is combined with a novel algorithm to compute realistic estimate of building evacuation time. Additionally, this novel algorithm can be utilized in a real evacuation situation to help guide an individual to provide direction to the entire crowd to optimize the evacuation process. The overarching assumption in this work is that everybody listen and follows the guidance provided. This assumption is necessary to simplify the problem and introduce a working solution. In future work, this overarching assumption will be relaxed and the effect of different level of guidance acceptance will be studied. Also, scenarios with multiple guides can be investigated. Further, this realistic evacuation time estimator algorithm can be a starting point to compute evacuation strategy for a building under duress from an armed assailant. The nodes closest to the armed person should be given higher priority to evacuate and whether to barricade/move is a critical decision to preserve lives. These are some rewarding avenues to continue this work to improve the overall quality of existing evacuation strategies.

Chapter 6

Virtual reality setup to study factors affecting individuals' exit choice during emergency evacuation

6.1 Objective

So far we have mathematically modeled the opinion propagation in an egressing crowd (Chapter 2), developed an simulation environment to study the effect of different behavior on evacuation time, (Chapter3) and established an algorithm to compute realistic estimate of building evacuation time (Chapter 5). Finally, virtual reality offers an unique opportunity of safe environment where one can run different scenarios to elicit information from participants regarding emergency evacuation. This gives an ethical way for researchers to reproduce the same scenario for more than one participant without exposing anyone to life threatening situation. Recently, virtual reality has been used to study the effectiveness of exit signs in one research work [111]. In another work, virtual reality has been utilized to conduct training and evacuation drill for disaster preparedness in a virtual train station [112]. Moussaïd et al. [58] utilized a non-immersive virtual environment to study crowd behavior during emergency evacuation. The objective of this part of research is to reproduce an experiment conducted by Bode et al. [60] in an immersive virtual reality environment. This immersive replication

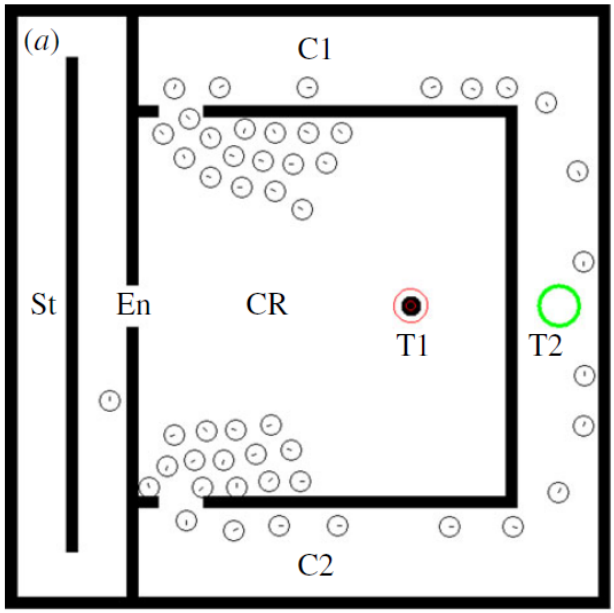


Figure 6.1: Illustration of the simulated environment utilized in Bode et al. experiment - Figure is sourced from [60]

of the existing body of work can help to understand the impact of immersion on the results obtained from the participants.

During emergency evacuation, individuals will have different directional information available to them. Individuals can visually see bright exit signs leading them towards safety. Individuals can follow other evacuees (herding). Or individuals can follow their usual path(familiarity/memory) to exit the building. How these different source of information play a role in an individuals exit choice during emergency evacuation is crucial to understand the evacuation process. In Bode et al. work [60], they concluded that exit sign played a dominant role in exit choice decision of their participants. But when conflicting source of information was available, the prominence of exit sign diminished. They presented a non-immersive top-down view of the environment as depicted in figure 6.1 to their participants. Since the participants could not feel like they were in a real-life emergency situation, it does not give an immersive experience.

To alleviate this problem, we have designed a similar environment with an immersive capability (virtual reality platform) to collect data from individuals. This can help to gain

an interesting insight into how immersion can produce a similar/different inference for a similar situation.

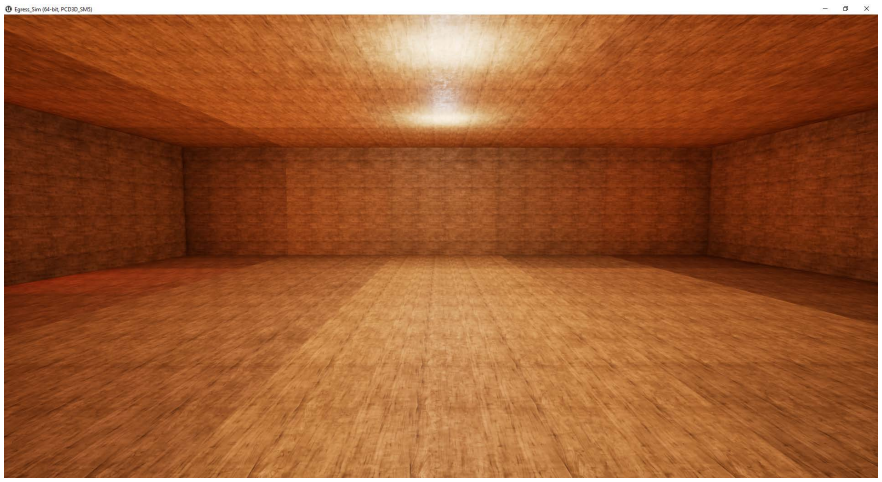
6.2 Virutal Reality Setup

For designing our immersive virtual reality environment Unreal game engine and HTC Vive platform were utilized. The individuals were first provided with an University IRB (Institutional Review Board) approved consent form. Participants were informed about the potential risks and benefits. After their consent, individuals were placed in a virtual room to learn navigating in a virtual environment. They had an immersive first person view (i.e.,) if the participant tilt their head down they would see their virtual lower body as depicted in the figure 6.2.

Next to establish a baseline of individuals quantitative aptitude and to elicit basic demographic information like their gender and age a paper survey similar to the one portrayed in figure 6.3 was provided to the subjects. Note that no personally identifiable information was collected in the survey and the quantitative section was timed. Typically a minute was provided to every participants to answer the quantitative questions. Individuals score in the quantitative section was correlated later with their performance in the emergency evacuation. As the final step in the virtual reality experimental data collection, each individual was shown a set of five scenarios. The participant started to evacuate from the back of the virtual room on hearing an audio alarm in all the five scenario. The virtual room will have two visible exits. Additionally, the simulated room will be populated with 40 virtual individuals who egress according to the preprogrammed scenario. The five scenarios are explained in detail in the following sections.

6.2.1 Exit sign scenario

In this one, there are again 40 virtual individuals evenly divided into two sides and they egress towards the visible exit on their respective side. The exit sign near one of the exits is lit and the evacuation process is recorded. According to published results of Bode et al.



(a)

(b)

Figure 6.2: (a) The tutorial room for participants to get acclimated to virtual environment (b) The view of their virtual lower body when they tilt their head down - immersive first person view

Demographic and Quantitative aptitude survey

Gender: ____ Male ____ Female ____Others ____ Prefer not to mention

Age : _____

Math Question 1:

1+10 = _____

Math Question 2:

10+51 = _____

Math Question 3:

145+642 = _____

Math Question 4:

100-10 = _____

Math Question 5:

2X4 = _____

Math Question 6:

13x15 = _____

Math Question 7:

51X78 = _____

Math Question 8:

12/4 = _____

Math Question 9:

143/11 = _____

Math Question 10:

5628/67 = _____

Figure 6.3: A sample of the demographic and quantitative aptitude survey provided to participants

Figure 6.4: A depiction of the first person view during the evacuation in an exit sign lit scenario

[60], we expect the human controlled virtual agent to more or less chose the exit with the exit sign lit. The figure 6.4 describes the exit sign lit scenario.

6.2.2 Crowd Scenario

In this one, there are again 40 virtual individuals evenly divided into two sides. But, when the evacuation starts, all of them pile up towards one of the exits. This scenario helps to elicit how much of a herd mentality participants possess. If the virtual human controlled individual move towards the exit where the crowd is piling up, it confirms the herding mentality. The figure 6.5 illustrates the crowd scenario.

6.2.3 Exit and crowd reinforcing scenario

In this scenario, there is a exit sign lit and the entire virtual crowd is moving towards the exit with that sign. In this one, as the name suggests, the 40 virtual individuals move towards one of the available exit en masse. Also that exit will have a lit exit sign. In this scenario, we expect the individual controlled virtual agent to move with the crowd towards the exit with the lit sign. The figure 6.6 illustrates this scenario.

Figure 6.5: A depiction of the first person view during the evacuation in a crowd scenario

Figure 6.6: A depiction of the first person view during the evacuation in an exit sign lit along with reinforcing crowd scenario

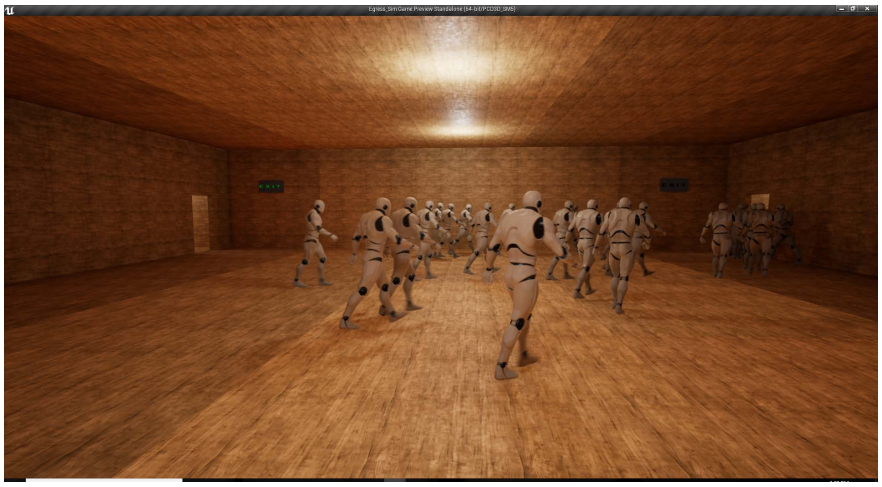


Figure 6.7: A depiction of the first person view during the evacuation in an exit sign lit along with opposing crowd scenario

6.2.4 Exit and crowd opposing scenario

In this scenario, there is a exit sign lit and the entire virtual crowd is moving towards the exit without the lit sign (i.e,) the 40 virtual individuals move towards the exit without the lit exit sign. Here, the objective is to capture the effect of conflicting directional information. In this scenario, we expect the individuals moving towards lit exit sign to reduce from the baseline established in the exit sign scenario. The figure 6.7 illustrates this scenario.

6.2.5 Control group scenario

In the control group scenario there are 40 virtual individuals moving towards the visible exit on their side of the room. The group of 40 in evenly divided into 20 each on either side. There is no lit exit signs and this control group is utilized to capture any bias the participants have towards a particular side/exit. The expected outcome of the control group is the participants are equally likely to egress towards either of the two available sides/exits. The figure 6.8 depicts the control scenario.

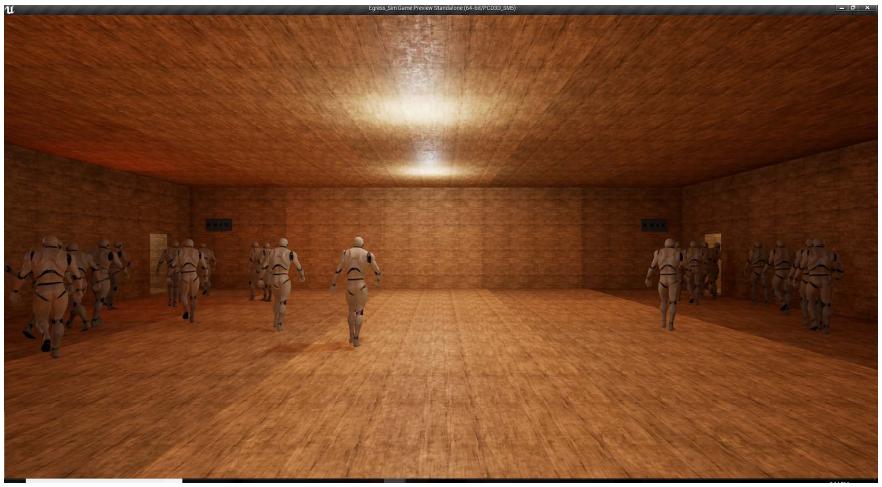


Figure 6.8: A depiction of the first person view during the evacuation in a control group scenario

6.3 Results and Discussion

A sample data collection with 11 individuals was conducted. There were 2 female participants and the rest were male participants. The average score on the quantitative section is 8.18 with a standard deviation of 0.60. This is expected since the study population mainly consisted of graduate students. Since, everyone scored about 8 (without much deviation) the analysis based on quantitative score with different exit choice can not be performed. The average age of the study population is 27.7 years old with a standard deviation of 4.8.

Figure 6.9(e) shows that there is no statistically significant preference among the participants towards any of the two available exits (i.e,) there is no existing bias towards a particular exit. Comparing to existing notion established in [60], the exit sign only scenario did not elicit a strong attraction towards the exit with the lit exit sign (Fig. 6.9(a)). The study population did not exhibit any preference towards the lit exit sign other than that can be explained by chance. This is also true (Fig. 6.9(b)) when there are no exit sign/and the crowd is moving towards one of the exits en mass (scenario 2). This was also the case in the scenario with the crowd moving towards the lit exit sign (scenario 3) (Fig. 6.9(c)). This result can be due to the fact that the individuals were informed prior to the virtual data collection that there were two available exits.

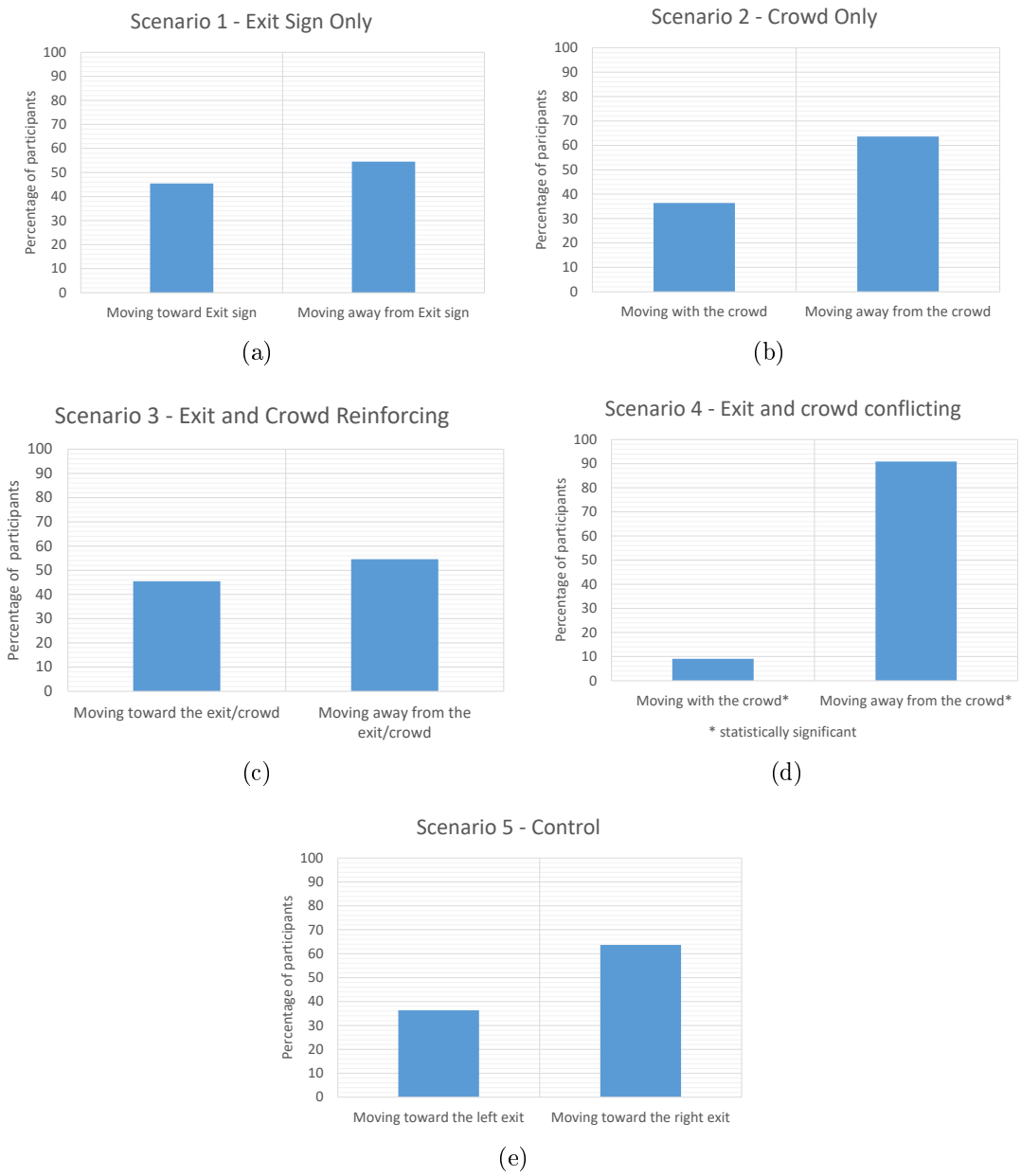


Figure 6.9: Bar graph depicting percentage of people out of 11 total participants (a) following/not following the exit sign - Scenario 1, (b) following/not following the crowd - Scenario 2, (c) following/not following the crowd and the exit sign - Scenario 3, (d) following/not following the crowd in crowd and exit sign conflicting scenario, and (e) preferred exit in control scenario - Scenario 5

Finally, in the situation with exit sign lit on one side and the crowd moving towards the exit on the other side (scenario 4), there was statistically significant exhibition of aversion to the crowd (Fig. 6.9(d)). In this virtual reality based immersive experience, the participants can feel the congestion when they move with the crowd. This is not the case with the top down view of environment in Bode et al. [60] work. There was no sense of being physically congested in their computer monitor based setup. This feeling of being crowded in the virtual environment along with the knowledge of two available exits, elicited the crowd aversion behavior from the participants. From this set of results, the herding parameter corresponding to the model established in chapter 3 was computed utilizing the following equation along the lines of the spatially bounded confidence opinion sharing model equation 3.2.

$$\begin{pmatrix} \text{\% of individuals moving with the crowd} \\ \text{\% of individuals moving away from crowd} \end{pmatrix} = (1-\mu) \times \text{value matrix of self+} \\ \mu \times \text{average value matrix of the crowd} + \epsilon \tag{6.1}$$

where μ is the herding parameter and the individual's give μ % weightage to the crowd's opinion. Here, the self value matrix was structured such that it reflect the choice of avoiding the crowd.

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{6.2}$$

The crowd average value matrix reflects that the entire crowd is moving towards one of the exits.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{6.3}$$

Finally, the ϵ is to account for unaccounted parameters playing a role in participants' exit choice. Substituting the values from the collected data, we get

$$\begin{pmatrix} 0.091 \\ 0.909 \end{pmatrix} = (1-\mu) \times \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \mu \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \epsilon \tag{6.4}$$

Table 6.1: Herding Parameter (μ) values from different scenarios utilized in the VR based data collection

Scenario	Estimated μ	Comments
1	N/A	There is no crowd bias to extrapolate herding parameter
2	0.36	In presence of just the crowd, 36% of participants followed the crowd
3	N/A	The effect of exit sign and the crowd can not be distinguished from each other
4	0.09	In presence of crowd and exit sign conflicting each other only 9% of participants followed the crowd. Their primary motive seems to be avoiding congestion at the exit
5	[0 , 1]	Any value of μ will sufficiently explain the observed characteristics since the crowd is evenly divided between the 2 exits.

Solving equation 6.3, the herding parameter μ value is 0.09 which implies that the participants were giving approximately 9% weightage to the crowds' opinion/exit choice. Similarly for scenario 2, if the herding parameter is computed, the value will be 0.36. This implies that approximately 36% weightage is given by individuals to crowds opinion in absence of any other information source. These results are presented in the table **??** More complex building structure and different virtual experiment design can help to extract other parameters of the model like the impatience factor, decision timing, etc.

From this virtual reality data collection, some existing notion about directional information stands invalidated. It will be interesting to conduct a more extensive data collection to validate the results obtained from this pilot study. From this data collection, it is observed that the participants have a strategy of avoiding the crowd to eliminate potential congestion and subsequent increase in evacuation time. Also, the participants did not pay much attention to exit sign as predicted by previous works. This can be due to the fact that the participants knew about the availability of two exits. When conflicting information was provided, the participants chose to follow the exit sign in order to avoid the crowd.

6.4 Summary

According to existing data and inference in Bode et al. [60] work, the exit sign should be the most prominent directional information utilized by the participants. But, when conflicting information is provided by the crowd, the proportion of participants moving towards the lit exit sign is expected to diminish. Thus, even though the exit sign is the most prominent directional information, in presence of another conflicting source of directional information, the reliance on exit sign is expected to diminish. The data collected in the virtual immersive environment is contrary to this existing notion. The participants followed the exit sign more when conflicting information through crowd movement was present. The participants main strategy was to avoid the crowd and the subsequent congestion at the exit. This new insight can lead to designing a reliable guidance delivery mechanism through different directional information source to help individuals in a crowd to move towards safety in a real life emergency situation.

Chapter 7

Conclusion

7.1 Research Overview

The overarching goal of this research work is to further our understanding of the emergency evacuation and improve existing strategies for evacuation. The research goals are four fold. The first is to establish a mathematical model of the opinion sharing among individuals in an egressing crowd. The second goal is to develop a simulation model from scratches which can mimic various real-life observable phenomena like herding, impatience in a mathematically based simulation and conduct an extensive parametric study. The third, developing an algorithm for estimating a realistic evacuation time for a given building structure by adapting existing algorithms. And the final objective is to collect high fidelity data from human participants to study the effect of different directional information sources on individuals' exit choice. The common underlying theme or scientific objective across the above-mentioned goals is to better understanding of mechanisms involved in an emergency evacuation situation and providing practical solutions to improve the existing evacuation strategy.

7.2 Contributions

Objective 1 *(addressing gap 1)*: **Mathematical modeling and numerical validation of decision sharing.** In the first phase of the project, the combined movement and decision sharing model will be defined using stochastic differential equation. The time evolution of

the opinion (decision) propagation through the crowd will be mathematically derived from the stochastic equation. A corresponding simulation model will be developed and the effect of leaders or strong opinionated individuals on the overall distribution of the crowd at the available exits will be investigated.

- Derived a mathematical model for the opinion propagation through an egressing crowd (Chapter 2)
- Developed a Monte-Carlo simulation with presence of leaders or strong opinionated individuals and studied the effect of them on the overall distribution of the corwd at the available exits (Chapter 2)

Objective 2 *(addressing gap 2)* **: A Markov decision process based decision theoretic model with spatially bounded opinion sharing framework.** In the second stage of the project, a Markov decision process based decision model will be defined. The model will take into account the collision avoidance behavior of individuals, impatience exhibited by crowd at bottlenecks, re-evaluation of current route choice by evacuees at regular intervals, and herding behavior through a spatially bounded opinion sharing framework. A thorough parametric study of the factors affecting the overall efficiency of the evacuation process will be conducted.

- A Markovian model was proposed with mentioned behaviors mathematically accounted for in the hybrid model (Chapter 3)
- A thorough parametric study of the developed hybrid model is performed and the results were presented and discussed (Chapter 3)
- A modified algorithm to extract the reward function (the primary descriptor of the underlying Markov decision model) from a demonstration is presented (Chapter 4)
- The modified learning algorithm is tested with a toy robot navigation problem in a $2D$ environment to verify the algorithm (correctness and time scaling property) (Chapter 4)

Objective 3 *(addressing gap 3)* **: A dynamic guidance algorithm.** The lessons learned through implementation of objective 2 will be accounted for in a dynamic guidance algorithm. The guidance system will monitor current status of evacuation and update the guidance provided to the evacuees to improve the overall efficiency of the process. This will result in a realistic estimate of evacuation time for a given building structure. Also, a guide can be given a set of active instructions to navigate through the building to optimize the evacuation time of the evacuees.

- Implemented a existing near optimal heuristic algorithm to compute the optimal evacuation strategy for a given building structure (Chapter 5)
- Developed a niche algorithm that can take the optimal strategy along with the normal route preference of individuals and provide solution for a guide to reduce the gap between the normal and the optimal strategy (Chapter 5)

Objective 4 *(addressing gap 3)* **: A virtual reality set up to validate the factors affecting individuals' exit choice.** Having developed a guidance model and a set of instruction for a guide to improve the emergency evacuation procedure, the final task will involve developing an immersive virtual reality set up to test and validate the factors affecting individuals' decision. Virtual reality platform serves as an excellent tool to test out emergency evacuation scenarios since it does not put any individual through a real-life threatening situation. Nevertheless, virtual reality platform will help to elicit realistic information when compared to written/oral survey techniques about individual's choices during emergency evacuation. This can lead to better delivery of guidance to the evacuating crowd, thus improving the overall evacuation of the crowd.

- Utilized Unreal game engine and HTC Vice headset to create a virtual reality immersive environment (Chapter 6)
- Collected preliminary data and analyzed the participants behavior. The inferences were interesting and future data collection and analysis can look into a broader scope

of emergency evacuation related information. Further, an estimate for the herding parameter introduced in Chapter 3 was computed. (Chapter 6)

7.3 Future Direction

There are multiple research directions that can be pursued utilizing this body of research work. Some of the interesting potential avenues are expanding the mathematical modeling of opinion sharing into a 2D space where one dimension is the opinion space (Chapter 2) and the other dimension is the movement space. It would be an intriguing mathematical problem involving stochastic differential equations. Another avenue of research is to take the realistic building time evacuation estimator algorithm(Chapter 5) and modify it to include situations like active shooter scenario and come up with an elegant solution to prioritize evacuating individuals near high risk zone. Also, the question of whether it is safe to evacuate or barricade oneself to minimize risk when answered can lead to improved safety of individuals. Additionally, the rudimentary virtual reality environment (Chapter 6) can be further developed and include multi-player capability to acquire model parameters for the simulation engine developed in Chapter 3. Also, virtual reality environment can be utilized to learn the underlying intrinsic reward function (Chapter 4).

Bibliography

[1] Mark Follman, Gavin Aronsen, and DeAnna Pan. *US Mass Shootings, 1982-2017: Data from Mother Jones' Investigation*. Web Page. 2016. URL: http://www.motherjones.com/politics/2012/12/mass-shootings-mother-jones-full-data.

[2] National Fire Protection Association. *Fire Statistics*. Web Page. 2016. URL: http://www.nfpa.org/news-and-research/fire-statistics-and-reports/fire-statistics.

[3] RL Hughes. "The flow of large crowds of pedestrians". In: *Mathematics and Computers in Simulation* 53.4 (2000), pp. 367–370.

[4] Roger L Hughes. "A continuum theory for the flow of pedestrians". In: *Transportation Research Part B: Methodological* 36.6 (2002), pp. 507–535.

[5] Roger L Hughes. "The flow of human crowds". In: *Annual review of fluid mechanics* 35.1 (2003), pp. 169–182. DOI: 10.1146/annurev.fluid.35.101101.161136. URL: https://doi.org/10.1146/annurev.fluid.35.101101.161136.

[6] Rinaldo M Colombo and Massimiliano D Rosini. "Pedestrian flows and non-classical shocks". In: *Mathematical Methods in the Applied Sciences* 28.13 (2005), pp. 1553–1567.

[7] Jan Dijkstra, Harry JP Timmermans, and AJ Jessurun. "A multi-agent cellular automata system for visualising simulated pedestrian activity". In: *Theory and Practical Issues on Cellular Automata*. Springer, 2001, pp. 29–36.

[8] Victor J. Blue and Jeffrey L. Adler. "Cellular automata microsimulation for modeling bi-directional pedestrian walkways". In: *Transportation Research Part B: Methodological* 35.3 (2001), pp. 293–312. ISSN: 0191-2615. DOI: http://dx.doi.org/10.1016/S0191-2615(99)00052-1.

[9] Carsten Burstedde et al. "Simulation of pedestrian dynamics using a two-dimensional cellular automaton". In: *Physica A: Statistical Mechanics and its Applications* 295.3 (2001), pp. 507–525.

[10] Ansgar Kirchner and Andreas Schadschneider. "Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics". In: *Physica A: Statistical Mechanics and its Applications* 312.1 (2002), pp. 260–276.

[11] Ansgar Kirchner et al. "Discretization effects and the influence of walking speed in cellular automata models for pedestrian dynamics". In: *Journal of Statistical Mechanics: Theory and Experiment* 2004.10 (2004), P10011.

[12] Katsuhiro Nishinari et al. "Modelling of self-driven particles: Foraging ants and pedestrians". In: *Physica A: Statistical Mechanics and its Applications* 372.1 (2006), pp. 132–141. ISSN: 0378-4371. DOI: http://dx.doi.org/10.1016/j.physa.2006.05.016.

[13] Ansgar Kirchner et al. "Simulation of competitive egress behavior: comparison with aircraft evacuation data". In: *Physica A: Statistical Mechanics and its Applications* 324.3 (2003), pp. 689–697.

[14] Hubert Ludwig Kluepfel. "A cellular automaton model for crowd movement and egress simulation". PhD thesis. Universität Duisburg-Essen, Fakultät für Physik, 2003.

[15] Colin M Henein and Tony White. "Macroscopic effects of microscopic forces between agents in crowd models". In: *Physica A: statistical mechanics and its applications* 373 (2007), pp. 694–712.

[16] A Varas et al. "Cellular automaton model for evacuation process with obstacles". In: *Physica A: Statistical Mechanics and its Applications* 382.2 (2007), pp. 631–642.

[17] Nirajan Shiwakoti et al. "Animal dynamics based approach for modeling pedestrian crowd egress under panic conditions". In: *Transportation Research Part B: Methodological* 45.9 (2011), pp. 1433–1449. ISSN: 0191-2615. DOI: http://dx.doi.org/10.1016/j.trb.2011.05.016.

[18] Michael Seitz, Gerta Köster, and Alexander Pfaffinger. "Pedestrian group behavior in a cellular automaton". In: *Pedestrian and Evacuation Dynamics 2012*. Springer, 2014, pp. 807–814.

[19] R. Alizadeh. "A dynamic cellular automaton model for evacuation process with obstacles". In: *Safety Science* 49.2 (2011), pp. 315–323. ISSN: 0925-7535. DOI: http://dx.doi.org/10.1016/j.ssci.2010.09.006. URL: http://www.sciencedirect.com/science/article/pii/S0925753510002262.

[20] Dirk Helbing and Peter Molnar. "Social force model for pedestrian dynamics". In: *Physical review E* 51.5 (1995), p. 4282.

[21] DR Parisi and CO Dorso. "Microscopic dynamics of pedestrian evacuation". In: *Physica A: Statistical Mechanics and its Applications* 354 (2005), pp. 606–618.

[22] Daniel R Parisi and Claudio O Dorso. "Morphological and dynamical aspects of the room evacuation process". In: *Physica A: Statistical Mechanics and its Applications* 385.1 (2007), pp. 343–355.

[23] M. Zhou et al. "Modeling of Crowd Evacuation With Assailants via a Fuzzy Logic Approach". In: *Ieee Transactions on Intelligent Transportation Systems* 17.9 (2016), pp. 2395–2407. ISSN: 1524-9050. DOI: 10.1109/tits.2016.2521783.

[24] Dirk Helbing. "Traffic and related self-driven many-particle systems". In: *Reviews of modern physics* 73.4 (2001), p. 1067.

[25] Dirk Helbing and Anders Johansson. "Pedestrian, Crowd and Evacuation Dynamics". In: *Extreme Environmental Events: Complexity in Forecasting and Early Warning*. New York, NY: Springer New York, 2011, pp. 697–716. ISBN: 978-1-4419-7695-6. DOI: 10.1007/978-1-4419-7695-6_37.

[26] Dirk Helbing et al. "Lattice gas simulation of experimentally studied evacuation dynamics". In: *Physical review E* 67.6 (2003), p. 067101.

[27] Kouhei Takimoto and Takashi Nagatani. "Spatio-temporal distribution of escape time in evacuation process". In: *Physica A: Statistical Mechanics and its Applications* 320 (2003), pp. 611–621.

[28] Weiguo Song et al. "Simulation of evacuation processes using a multi-grid model for pedestrian dynamics". In: *Physica A: Statistical Mechanics and its Applications* 363.2 (2006), pp. 492–500.

[29] RY Guo and Hai-Jun Huang. "A mobile lattice gas model for simulating pedestrian evacuation". In: *Physica A: Statistical Mechanics and its Applications* 387.2 (2008), pp. 580–586.

[30] Paolo Lino, Guido Maione, and Bruno Maione. "Modeling and simulation of crowd egress dynamics in a discrete event environment". In: *Control Applications,(CCA) & Intelligent Control,(ISIC), 2009 IEEE*. IEEE. 2009, pp. 843–848.

[31] Harmeet Singh et al. "Modelling subgroup behaviour in crowd dynamics DEM simulation". In: *Applied Mathematical Modelling* 33.12 (2009), pp. 4408–4423.

[32] Siu Ming Lo et al. "A game theory based exit selection model for evacuation". In: *Fire Safety Journal* 41.5 (2006), pp. 364–369.

[33] Michel Bierlaire, Gianluca Antonini, and Mats Weber. *Behavioral dynamics for pedestrians*. Tech. rep. IEEE, 2003.

[34] Gianluca Antonini, Michel Bierlaire, and Mats Weber. "Discrete choice models of pedestrian walking behavior". In: *Transportation Research Part B: Methodological* 40.8 (2006), pp. 667–687. ISSN: 0191-2615. DOI: http://dx.doi.org/10.1016/j.trb.2005.09.006.

[35] Th Robin et al. "Specification, estimation and validation of a pedestrian walking behavior model". In: *Transportation Research Part B: Methodological* 43.1 (2009), pp. 36–56. ISSN: 0191-2615. DOI: http://dx.doi.org/10.1016/j.trb.2008.06.010.

[36] Serge P Hoogendoorn and Piet HL Bovy. "Pedestrian route-choice and activity scheduling theory and models". In: *Transportation Research Part B: Methodological* 38.2 (2004), pp. 169–190. ISSN: 0191-2615. DOI: http://dx.doi.org/10.1016/S0191-2615(03)00007-9.

[37] Ruggiero Lovreglio et al. "A discrete choice model based on random utilities for exit choice in emergency evacuations". In: *Safety science* 62 (2014), pp. 418–426.

[38] Ruggiero Lovreglio et al. "The role of herding behaviour in exit choice during evacuation". In: *Procedia-Social and Behavioral Sciences* 160 (2014), pp. 390–399.

[39] Xiaoshan Pan et al. "A computational framework to simulate human and social behaviors for egress analysis". In: *Proceedings of the Joint International Conference on Computing and Decision Making in Civil and Building Engineering.* 2006, pp. 1206–1215.

[40] Xiaoshan Pan et al. "A multi-agent based framework for the simulation of human and social behaviors during emergency evacuations". In: *Ai & Society* 22.2 (2007), pp. 113–132.

[41] Xiaoshan Pan. "Computational modeling of human and social behaviors for emergency egress analysis". PhD thesis. Stanford University, 2006.

[42] Mei Ling Chu, Xiaoshan Pan, and Kincho Law. "Incorporating social behaviors in egress simulation". In: *Proceedings of 2011 Computing in Civil Engineering Workshop.* 2011, pp. 19–22.

[43] Nuria Pelechano and Norman I Badler. "Modeling crowd and trained leader behavior during building evacuation". In: *IEEE computer graphics and applications* 26.6 (2006).

[44] Timo Korhonen and Simo Heliövaara. "Fds+ evac: herding behavior and exit selection". In: *Fire Safety Science* 10 (2011), pp. 723–734.

[45] Guylène Proulx. *Understanding human behaviour in stressful situations.* 2002.

[46] Samiul Hasan and Satish V Ukkusuri. "Social contagion process in informal warning networks to understand evacuation timing behavior". In: *Journal of Public Health Management and Practice* 19 (2013), S68–S69.

[47] Samiul Hasan and Satish V Ukkusuri. "A threshold model of social contagion process for evacuation decision making". In: *Transportation research part B: methodological* 45.10 (2011), pp. 1590–1605. ISSN: 0191-2615. DOI: http://dx.doi.org/10.1016/j.trb.2011.07.008.

[48] Robert L Goldstone and Marco A Janssen. "Computational models of collective behavior". In: *Trends in cognitive sciences* 9.9 (2005), pp. 424–430.

[49] Kristjan Bergey, Kevin Spieser, and Daniel E Davison. "The psychological dynamics of students in a classroom: Modeling and control strategies based on suggestibility theory". In: *Control Applications, 2007. CCA 2007. IEEE International Conference on*. IEEE. 2007, pp. 658–663.

[50] Kevin Spieser and DE Davison. "Stabilizing the psychological dynamics of people in a queue". In: *American Control Conference, 2008*. IEEE. 2008, pp. 4173–4178.

[51] Kevin Spieser and Daniel E Davison. "Multi-agent stabilisation of the psychological dynamics of one-dimensional crowds". In: *Automatica* 45.3 (2009), pp. 657–664.

[52] Kevin Spieser. "Stabilizing the psychological dynamics of people in a crowd". PhD thesis. University of Waterloo, 2008.

[53] Gustave Le Bon. *The crowd: A study of the popular mind*. Fischer, 1897.

[54] John Drury et al. "Cooperation versus competition in a mass emergency evacuation: A new laboratory simulation and a new theoretical model". In: *Behavior research methods* 41.3 (2009), pp. 957–970.

[55] Alexandre Nicolas, SebastiÃąn Bouzat, and Marcelo N. Kuperman. "Pedestrian flows through a narrow doorway: Effect of individual behaviours on the global flow and microscopic dynamics". In: *Transportation Research Part B: Methodological* 99 (2017), pp. 30–43. ISSN: 0191-2615. DOI: http://dx.doi.org/10.1016/j.trb.2017.01.008.

[56] Mehdi Moussaïd et al. "Experimental study of the behavioural mechanisms underlying self-organization in human crowds". In: *Proceedings of the Royal Society of London B: Biological Sciences* 276.1668 (2009), pp. 2755–2762.

[57] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. "How simple rules determine pedestrian behavior and crowd disasters". In: *Proceedings of the National Academy of Sciences* 108.17 (2011), pp. 6884–6888.

[58] Mehdi Moussaïd et al. "Crowd behaviour during high-stress evacuations in an immersive virtual environment". In: *Journal of The Royal Society Interface* 13.122 (2016). DOI: 10.1098/rsif.2016.0414.

[59] Nikolai WF Bode and Edward A Codling. "Human exit route choice in virtual crowd evacuations". In: *Animal Behaviour* 86.2 (2013), pp. 347–358.

[60] Nikolai WF Bode, Armel U Kemloh Wagoum, and Edward A Codling. "Human responses to multiple sources of directional information in virtual crowd evacuations". In: *Journal of The Royal Society Interface* 11.91 (2014), p. 20130904.

[61] Yuan Gao et al. "Fire evacuation model with confidence intervals". In: *Automation Science and Engineering (CASE), 2011 IEEE Conference on*. IEEE. 2011, pp. 731–736.

[62] Peng Wang et al. "Modeling and optimization of crowd guidance for building emergency evacuation". In: *Automation Science and Engineering, 2008. CASE 2008. IEEE International Conference on*. IEEE. 2008, pp. 328–334.

[63] ED Kuligowski. "Review of 28 egress models". In: *NIST SP* 1032 (2004).

[64] Xiaoping Zheng, Tingkuan Zhong, and Mengting Liu. "Modeling crowd evacuation of a building based on seven methodological approaches". In: *Building and Environment* 44.3 (2009), pp. 437–445.

[65] Dorine C. Duives, Winnie Daamen, and Serge P. Hoogendoorn. "State-of-the-art crowd motion simulation models". In: *Transportation Research Part C: Emerging Technologies* 37 (2013), pp. 193–209. ISSN: 0968-090X. DOI: http://dx.doi.org/10.1016/j.trc.2013.02.005.

[66] Jason D Averill. "Five grand challenges in pedestrian and evacuation dynamics". In: *Pedestrian and Evacuation Dynamics*. Springer, 2011, pp. 1–11.

[67] Subhadeep Chakraborty. "Analytical methods to investigate the effects of external influence on socio-cultural opinion evolution". In: *Social Computing, Behavioral-Cultural Modeling and Prediction*. Springer, 2013, pp. 386–393.

[68] Aravinda R Srinivasan and Subhadeep Chakraborty. "Effect of network topology on the controllability of voter model dynamics using biased nodes". In: *American Control Conference (ACC), 2014*. IEEE. 2014, pp. 2096–2101.

[69] E Wong and JB Thomas. "On polynomial expansions of second-order distributions". In: *Journal of the Society for Industrial and Applied Mathematics* 10.3 (1962), pp. 507–516.

[70] Aravinda Ramakrishnan Srinivasan, Farshad Salimi Naneh Karan, and Subhadeep Chakraborty. "Pedestrian dynamics with explicit sharing of exit choice during egress through a long corridor". In: *Physica A: Statistical Mechanics and its Applications* (2016). ISSN: 0378-4371. DOI: http://dx.doi.org/10.1016/j.physa.2016.11.118.

[71] W Challenger, WC Clegg, and AM Robinson. "Understanding crowd behaviours: Guidance and lessons identified". In: *UK Cabinet Office* (2009).

[72] Martin L Puterman. *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons, 2014.

[73] Richard Bellman. "A Markovian Decision Process". In: *Indiana Univ. Math. J.* 6 (4 1957), pp. 679–684. ISSN: 0022-2518.

[74] Neville Q Verlander and Benjamin G Heydecker. "Pedestrian route choice: An empirical study". In: PTRC Education and Research Services Ltd. 1997.

[75] Dirk Helbing, Illés Farkas, and Tamas Vicsek. "Simulating dynamical features of escape panic". In: *Nature* 407.6803 (2000), pp. 487–490. DOI: http://dx.doi.org/10.1038/35035023.

[76] Farshad Salimi Naneh Karan, Aravinda Ramakrishnan Srinivasan, and Subhadeep Chakraborty. "Modeling and numerical simulations of the influenced Sznajd model". In: *Physical Review E* 96.2 (2017), p. 022310.

[77] Farshad Salimi Naneh Karan and Subhadeep Chakraborty. "A Parametric Study of Opinion Progression in a Divided Society". In: *International Conference on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*. Springer. 2017, pp. 182–192.

[78] Farshad Salimi Naneh Karan and Subhadeep Chakraborty. "Detecting behavioral anomaly in social networks using symbolic dynamic filtering". In: *ASME 2015 Dynamic Systems and Control Conference*. American Society of Mechanical Engineers. 2015, V003T37A001–V003T37A001.

[79] Farshad Salimi Naneh Karan and Subhadeep Chakraborty. "Dynamics of a repulsive voter model". In: *IEEE Transactions on Computational Social Systems* 3.1 (2016), pp. 13–22.

[80] Jan Lorenz. "Continuous opinion dynamics under bounded confidence: A survey". In: *International Journal of Modern Physics C* 18.12 (2007), pp. 1819–1838. ISSN: 0129-1831. DOI: http://dx.doi.org/10.1142/S0129183107011789.

[81] Guillaume Deffuant et al. "Mixing beliefs among interacting agents". In: *Advances in Complex Systems* 3.01n04 (2000), pp. 87–98. ISSN: 0219-5259.

[82] Rainer Hegselmann and Ulrich Krause. "Opinion dynamics and bounded confidence models, analysis, and simulation". In: *Journal of Artificial Societies and Social Simulation* 5.3 (2002).

[83] Farshad Salimi Naneh Karan and Subhadeep Chakraborty. "Effect of Zealots on the opinion dynamics of rational agents with bounded confidence". In: *Acta Physica Polonica B* 49.1 (2018).

[84] Frank J Massey Jr. "The Kolmogorov-Smirnov test for goodness of fit". In: *Journal of the American statistical Association* 46.253 (1951), pp. 68–78.

[85] Andrew Y Ng, Stuart J Russell, et al. "Algorithms for inverse reinforcement learning." In: *Icml*. 2000, pp. 663–670.

[86] Brenna D Argall et al. "A survey of robot learning from demonstration". In: *Robotics and autonomous systems* 57.5 (2009), pp. 469–483.

[87] Alberto Maria Segre and Gerald DeJong. "Explanation-based manipulator learning: Acquisition of planning ability through observation". In: *Robotics and Automation. Proceedings. 1985 IEEE International Conference on*. Vol. 2. IEEE. 1985, pp. 555–560.

[88] Petar Kormushev, Sylvain Calinon, and Darwin G Caldwell. "Imitation learning of positional and force skills demonstrated via kinesthetic teaching and haptic input". In: *Advanced Robotics* 25.5 (2011), pp. 581–603.

[89] Paul Evrard et al. "Teaching physical collaborative tasks: Object-lifting case study with a humanoid". In: *Humanoid Robots, 2009. Humanoids 2009. 9th IEEE-RAS International Conference on*. IEEE. 2009, pp. 399–404.

[90] Seungsu Kim, Ashwini Shukla, and Aude Billard. "Catching objects in flight". In: *Robotics, IEEE Transactions on* 30.5 (2014), pp. 1049–1065.

[91] Seungsu Kim and Aude Billard. "Estimating the non-linear dynamics of free-flying objects". In: *Robotics and Autonomous Systems* 60.9 (2012), pp. 1108–1122.

[92] Sonia Chernova and Manuela Veloso. "Interactive policy learning through confidence-based autonomy". In: *Journal of Artificial Intelligence Research* 34.1 (2009), p. 1.

[93] Sonia Chernova and Manuela Veloso. "Multi-thresholded approach to demonstration selection for interactive robot learning". In: *Human-Robot Interaction (HRI), 2008 3rd ACM/IEEE International Conference on*. IEEE. 2008, pp. 225–232.

[94] Pieter Abbeel and Andrew Y Ng. "Apprenticeship learning via inverse reinforcement learning". In: *Proceedings of the twenty-first international conference on Machine learning*. ACM. 2004, p. 1.

[95] Pieter Abbeel and Andrew Y Ng. "Exploration and apprenticeship learning in reinforcement learning". In: *Proceedings of the 22nd international conference on Machine learning*. ACM. 2005, pp. 1–8.

[96] Beomjoon Kim and Joelle Pineau. "Socially adaptive path planning in human environments using inverse reinforcement learning". In: *International Journal of Social Robotics* (2015), pp. 1–16.

[97] B Kim and J Pineau. "Human-like navigation: Socially adaptive path planning in dynamic environments". In: *RSS 2013 Workshop on Inverse Optimal Control and Robotic Learning from Demonstration*. 2013.

[98] Peter Henry et al. "Learning to navigate through crowded environments". In: *Robotics and Automation (ICRA), 2010 IEEE International Conference on*. IEEE. 2010, pp. 981–986.

[99] Dizan Vasquez, Billy Okal, and Kai O Arras. "Inverse reinforcement learning algorithms and features for robot navigation in crowds: an experimental comparison". In: *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE. 2014, pp. 1341–1346.

[100] Kyriacos Shiarlis et al. "Inverse Reinforcement Learning from Failure". In: *RSS 2015: Proceedings of the 2015 Robotics: Science and Systems Conference, Workshop on Learning from Demonstration: Inverse Optimal Control, Reinforcement Learning, and Lifelong Learning*. 2015.

[101] Quoc Phong Nguyen, Bryan Kian Hsiang Low, and Patrick Jaillet. "Inverse Reinforcement Learning with Locally Consistent Reward Functions". In: *Advances in Neural Information Processing Systems*. 2015, pp. 1738–1746.

[102] Brian D Ziebart et al. "Maximum Entropy Inverse Reinforcement Learning." In: *AAAI*. 2008, pp. 1433–1438.

[103] Marc Deisenroth and Carl E Rasmussen. "PILCO: A model-based and data-efficient approach to policy search". In: *Proceedings of the 28th International Conference on machine learning (ICML-11)*. 2011, pp. 465–472.

[104] Hongqing Zhu. "Image representation using separable two-dimensional continuous and discrete orthogonal moments". In: *Pattern Recognition* 45.4 (2012), pp. 1540–1558.

[105] Richard Bellman. *The theory of dynamic programming*. Tech. rep. DTIC Document, 1954.

[106] Aravinda Ramakrishnan Srinivasan and Subhadeep Chakraborty. "Path planning with user route preference-A reward surface approximation approach using orthogonal Legendre polynomials". In: *Automation Science and Engineering (CASE), 2016 IEEE International Conference on*. IEEE. 2016, pp. 1100–1105.

[107] Thomas M Kisko and Richard L Francis. "EVACNET+: a computer program to determine optimal building evacuation plans". In: *Fire Safety Journal* 9.2 (1985), pp. 211–220.

[108] Peter Lin et al. "On the use of multi-stage time-varying quickest time approach for optimization of evacuation planning". In: *Fire Safety Journal* 43.4 (2008), pp. 282–290.

[109] M. Yusoff, J. Ariffin, and A. Mohamed. "Optimization approaches for macroscopic emergency evacuation planning: A survey". In: *2008 International Symposium on Information Technology*. Vol. 3. 2008, pp. 1–7. DOI: 10.1109/ITSIM.2008.4631982.

[110] Qingsong Lu, Yan Huang, and Shashi Shekhar. "Evacuation planning: a capacity constrained routing approach". In: *International Conference on Intelligence and Security Informatics*. Springer. 2003, pp. 111–125.

[111] Chieh-Hsin Tang, Wu-Tai Wu, and Ching-Yuan Lin. "Using virtual reality to determine how emergency signs facilitate way-finding". In: *Applied ergonomics* 40.4 (2009), pp. 722–730.

[112] Sharad Sharma et al. "Immersive virtual reality environment of a subway evacuation on a cloud for disaster preparedness and response training". In: *Computational Intelligence for Human-like Intelligence (CIHLI), 2014 IEEE Symposium on*. IEEE. 2014, pp. 1–6.

Vita

Aravinda Ramakrishnan was born in Kumbakonam, Tamil Nadu, India. He completed his higher secondary education at A.R.R. Matriculation Higher Secondary School, Kumbakonam as one of the school toppers. He graduated with his undergraduate degree in Electronics and Communication Engineering from SASTRA University, Thanjavur, India in May 2013. As part of his senior year capstone project, he worked as a research intern at Indian Institute of Technology, Powai, Mumbai. He moved to Knoxville, Tennessee to pursue his doctoral degree in Mechanical Engineering with control systems concentration under the guidance of Prof. Chakraborty in August 2013. As a graduate research assistant under Prof. Chakraborty, he worked on several interesting projects and dissertation was focused on mathematical modeling of emergency evacuation process and improving existing evacuation guidance algorithm. His research work has resulted in a couple of journal publications and 5 conference publications/presentations as of April 2018. Other than being a meticulous researcher, Aravinda participates in many volunteer activities. He has been an active volunteer with American Red Cross, Tennessee Science Bowl, Tennessee Science Olympiad to name a few. He takes pride in spreading STEM interest amongst school students by participating in departmental outreach activities regularly. Additionally, he has served as an active executive member with TechCarniVOL, a student led organization that promotes STEM interest in high-schoolers and freshmen by conducting intriguing competitions for them to participate and win cash prizes. For recreation, he loves to go whitewater kayaking, mountain biking and play basketball.