Investigation of Pedestrian Movement in Groups and in High-density Bottlenecks Using Discrete Choice Modelling Framework

by

Zohreh Rashedi-Ashrafi

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Master of Applied Science

in

Civil Engineering

Carleton University
Ottawa, Ontario

To my parents, Alireza and Parvaneh Rashedi

Abstract

Walking as a non-motorized mode of transport is an essential component of sustainable environment. To increase the share of walking in transportation and accordingly to plan and manage pedestrian areas, a deep understanding of pedestrian movement behaviour is required. Developing a model that can reproduce pedestrian behaviour can be used as a tool for assessing the existing and designing new pedestrian spaces. The aim of this thesis is to provide walking behaviour models based on discrete choice framework first to study the behaviour of pedestrians walking in groups and second to investigate pedestrian movements at bottlenecks under high density situations. Discrete choice models focus directly on behavioural aspects of pedestrian movement and model the choice behaviour of individuals when they have to select among a set of alternatives. The correlation between alternatives is captured by Cross-nested logit model in this thesis. Data has been extracted from real-world video recordings of pedestrian crowds in order to calibrate and validate walking behaviour models. In the first part of this thesis, group behaviour has been modeled based on the tendency of group members to maintain group unity. Maximum likelihood estimation has been used for the calibration purpose. The significance of social bonds between group members in presenting a realistic walking behaviour has been revealed in the estimation process. In the second part of the thesis, a simulation tool has been developed to calibrate and validate a model presenting pedestrian behaviour at high density bottlenecks. The validation results confirm the acceptable performance of walking models proposed in both parts of the thesis.

Acknowledgements

I would like to thank those people who supported me during my master's studies. First of all, I would like to express my sincere gratitude to Professor Karim Ismail, my supervisor, for all his guidance, support, and patience throughout this endeavor. I am very much thankful to him for the assistance and the encouragement he provided me during my graduate studies at Carleton University and deeply appreciate his belief in me. Moreover, I wish to thank the financial supporter of this research; the Natural Sciences and Engineering Research Council of Canada (NSERC). I also would like to thank Denis for helping me out through data extraction. Furthermore, I am very grateful to my best friends for helping me get through the difficult times with all the emotional support, entertainment, and caring they provided. Most importantly, none of this would have been possible without the love and patience of my family- my father, Alireza, my mother, Parvaneh, my sisters, Gelareh and Alaleh, and my brother, Hossein. They have been a constant source of love, support and encouragement during my postgraduate education. I would like to express my heart-felt gratitude to them.

Table of Contents

List of Tables

List of Illustrations

List of Acronyms

2-D	Two Dimensional
ANOVA	Analysis of Variance
ASC	Alternative Specific Constant
CA	Cellular Automata
CNL	Cross-Nested Logit
DCM	Discrete Choice Model
GEV	Generalized Extreme Value
HCM	Highway Capacity Manual
IIA	Independence from Irrelevant Alternatives
LOS	Level Of Service
LR	Likelihood Ratio
MLE	Maximum Likelihood Estimation
MNL	Multinomial Logit
NL	Nested Logit
RUM	Random Utility Model
SF	Social Force

Chapter One: Introduction

1.1 Research Motivation

The concept of sustainable environment has increasingly become so popular over the last years. Walking as the most sustainable mode of transport has attracted a lot of attention and as a result many efforts have been made to increase the share of walking in transport. The Reduction in fossil fuel usage and resource consumption is the main factor for considering walking a sustainable mode in comparison with other modes. In addition, walking has been well accepted to have many environmental, economic, and community benefits along with personal health benefits. The reduction of automobile usage and accordingly reduced pollutant and greenhouse gases such as carbon emission is one of the most significant benefits of choosing walking over other modes of transport. Increased urban livability, efficiency of land use including accessibility, reduced traffic congestion, transportation cost savings, and increased social interactions are some of its other benefits (Tolley 2003). Moreover, all the trips made by any means of transport inside cities begin and end with walking (Kashani Jou 2011). The majority of trips under a kilometre are also made by walking. It is the prevalent mode of transport in public spaces, commercial centres, and transport interchange facilities. Accordingly, planning for pedestrians and management of pedestrian areas is becoming more of interest among engineers and planners, which clearly requires a deep understanding of pedestrian movement behaviour.

The study of human dynamics and pedestrian behaviour is of great application to researchers in various fields such as transportation engineering, urban planning, architecture, social science, and computer games. In the context of transportation and urban planning, understanding and predicting pedestrian movement behaviour is significant from different aspects among which the most important ones are: design and operation assessment of pedestrian facilities (*e.g.,* public transport stations, shopping malls, and airports), crowd management, pedestrian safety improvement in indoor and outdoor environments, and dealing with congestion problems. For example, pedestrian walking time when accessing to or transferring from a transit facility plays an essential role in preparing the transit time-tables or in order to prevent excessive and unnecessary delays and also to avoid injuries and fatalities in crowded situations, a certain level of service has to be provided (Asano *et al.* 2010). The level of service of pedestrian facilities is determined based on the flow characteristics of pedestrian crowd. Therefore, it can be seen that in all the aforementioned scenarios, infrastructure designer and public transport planner clearly require to be familiar with pedestrian movement behaviour in order to optimize their designs (Hoogendoorn & Bovy 2001). Pedestrians do not behave similarly in different circumstances. Most of pedestrian issues and safety problems occur in critical conditions and in high density crowds. Therefore, evacuation process and panic situations (*e.g.,* crowded sport event evacuation or pedestrian fire evacuation) are other areas in which human movement behaviour is of great importance.

Developing a simulation model that can provide a realistic representation of pedestrian behaviour and their movements can be used as a tool to assess existing and to design new pedestrian spaces. Availability of visualization techniques in addition to the model itself offers strong evaluation tools for designers and planners. Therefore, a range of efficient techniques have been developed

by researchers over the years to describe and simulate the behaviour of individuals in various situations in order to study pedestrian traffic, and consequently, some well established models have been created in pedestrian context during recent years.

Modelling pedestrian behaviour is not a simple task. Contrary to driver behaviour, where vehicle movements are restricted to transport facilities (lanes) and are governed by traffic rules, walking is a complicated behaviour that cannot be represented by geometric of transport facilities. Pedestrians walk in a two-dimensional space and unlike drivers adjust their speed and direction at the same time. Many factors impact pedestrian's walking behaviour. It can be affected by human factors besides non-human factors and therefore vary from one person to another. All these reasons make pedestrians' behaviour difficult to predict. Moreover, modelling pedestrian behaviour involves a complex process of decision making which includes the pedestrian's perception of environment and the processing and assessing of the information received in order to make a decision (Papadimitriou *et al.* 2009; Hoogendoorn 2003; Camillen *et al.* 2009). In modelling process, it is necessary to identify accurate sets of behaviour that govern the motion of individuals in order to define how individuals respond to their environment. The model will be built based on these behavioural assumptions and proper mathematical framework. Environmental perception, maintaining personal space among individuals, collision avoidance, and orientation are all important factors, among others, that should be considered (Camillen *et al.* 2009).

Approaches that have been developed to model pedestrian behaviour can be categorized based on different criteria. However, the most common classification is macroscopic versus microscopic models. In recent years, more attention has been focused on microscopic models in which each pedestrian is modeled separately (as an agent) and individual pedestrian behaviour is

studied to investigate pedestrian traffic. These models deal with the behaviour of individuals including changing in agent's direction or speed, interactions between them, and navigation around obstacles. On the other hand, macroscopic models aim to model the pedestrian flow as a whole, mostly like a fluid. Similarity between fluid or gas dynamics and traffic flow is the basis of generating this type of models, using a set of differential equations. Pedestrian flow is described by means of aggregate traffic characteristics such as flow, average speed, and density (Shiwakoti & Nakatsuji 2005; Sahaleh *et al.* 2012). Interaction between pedestrians and also interaction between pedestrian and built-in environment considerably affect the macroscopic characteristics of pedestrian traffic. Therefore, investigating fundamental diagrams (the relationship between macroscopic flow characteristics) is an effective tool to study pedestrian dynamics (Ma *et al.* 2010).

1.2 Research Objectives and Scope

The focus of the research conducted in this thesis is on microscopic walking behaviour models based on discrete choice framework in order to study crowd dynamics and to reproduce the behaviour of pedestrians. Discrete choice analysis is used to model the choice behaviour of individuals when they have to select among a set of alternatives. Walking behaviour at operational level and in normal situations (*i.e.,* non-panic) is of interest in this study. Operational level of walking behaviour describes the instant decisions made by pedestrians while walking (Daamen 2004). Discrete choice models (DCM) are explained in detail in chapter 3 and different behavioural levels of pedestrian decision making are described later in Section 2.2.

The main objectives of this thesis can be divided into two parts. In the first part, the goal is to study group behaviour in discrete choice modelling framework. There are people walking in

groups as well as those walking individually in pedestrian crowds. The formations of groups and social bonds between group members have a major effect on pedestrians' behaviour. People in the same group maintain the group unity by performing some specific behaviours (Aveni 1977). Decisions made by each pedestrian in a group impact the behaviours of her group-mates as well. Therefore, group behaviour plays an essential role in understanding how pedestrians move in a crowd containing groups of pedestrians as well as individuals. Hence, the first objective of this research is to investigate the effect of grouping relationships on dynamics of crowd movements. In this research, this behaviour is presented as an extension to the discrete choice walking behaviour model proposed by Robin *et al.* (2009). Robin *et al.*'s work is described in section 3.2. As group behaviour has not been studied in DCM framework before, modelling this behaviour in a discrete choice framework is a novel contribution made in this thesis.

The objective of the second part of the research is to investigate the behaviour of pedestrian flows passing through bottlenecks in high density situations using discrete choice modelling framework. Dynamics of pedestrian crowds passing through bottlenecks have not been studied in DCM framework before. Since most of pedestrian issues and safety problems occur in critical conditions such as high density crowds, having a model that can provide a realistic representation of pedestrian behaviour in these conditions is greatly beneficial. This research aims to explore walking behaviour of unidirectional pedestrian flow when walking through a door in a relatively high density situation. In this study the discrete choice walking behaviour model proposed by Robin *et al.* (2009) is applied to simulate pedestrian behaviour in the aforementioned situations and the model's performance and the reliability of prediction is examined. This model was originally calibrated and validated on bi-directional pedestrian flow data, gathered in Japan and the Netherlands. However, pedestrians behave differently under

different circumstances and the factors affecting their behaviour may vary according to the environment and traffic conditions. Therefore, in this study, the pedestrian walking behaviour model proposed by Robin *et al.* (2009) is further developed and some changes are proposed to incorporate some aspects that have not already been addressed in the model in order to make it compatible with the new situation. The developed model is implemented in a simulation environment and the calibration and validation processes are carried out based on the macroscopic characteristics of crowd dynamics in this research.

Empirical data of human crowd is essential to study pedestrian behaviour. Pedestrian data under different circumstances are required to estimate pedestrian models or assess the predictive power of pedestrian models. However, it is difficult to collect accurate pedestrian-related data. In this research, the data was manually extracted from real-world video recordings of pedestrian crowds in order to calibrate and validate the walking behaviour model.

1.3 Thesis Outline

This chapter provides an introduction to this thesis. It presents the significance and the motivation of performing this research as well as the objectives of this research work.

The remaining chapters of this thesis are organized as follows: Chapter 2 presents a literature review regarding pedestrian walking behaviour. It provides some empirical data of pedestrian movement and gives an overview of self-organized patterns. The pedestrian hierarchy of decision making, interactions between pedestrians, and previous approaches to model walking behaviour are also explained in this chapter.

The focus of this research is on discrete choice modelling approach to model walking behaviour of pedestrians. Therefore, in chapter 3 general aspects of DCM framework along with the discrete choice pedestrian walking model developed by Robin *et al.* (2009) are described.

Chapter 4 studies group behaviour in discrete choice framework. The model specification for this behaviour, data collection, data extraction, data review process, calibration results, and validation results are provided in this chapter.

Chapter 5 aims to present a discrete choice walking behaviour model that provides acceptable functionality in high density situations for flows passing through bottlenecks. Data collection process and obtained calibration and validation results of the model are reported in this chapter.

Some conclusions for this research work and further research recommendations are given in chapter 6.

Chapter Two: Literature Review

This chapter provides an overview of previous studies in the context of pedestrian walking behaviour. Some empirical data of pedestrian dynamics and examples of self-organized patterns forming in pedestrian crowds are described in section 2.1. In section 2.2, a hierarchy of pedestrian behavioural levels is explained. Subsequently, some of the most important interactions of pedestrians are reviewed in section 2.3. Previous approaches developed to model the behaviour of pedestrians are discussed in section 2.4. Finally, summary of this chapter is given in section 2.5.

2.1 Empirical Study of Pedestrian Behaviour

Empirical data of pedestrian dynamics provides an insight into the characteristics and the walking behaviour of individuals and human crowds. Some previous findings regarding main macroscopic flow characteristics (flow, density, and speed) are briefly reviewed in this section. Subsequently, in the last part of the section, some self-organized patterns that form in crowds are discussed.

2.1.1 Fundamental Diagram

The aggregate behaviour of pedestrians is classically described by the macroscopic characteristics of traffic flow. These characteristics include density, flow, and speed of

pedestrian crowd. The relationship between them is represented by the traffic fundamental diagram, or as shorthand by the fundamental diagram. The capacity of pedestrian facility can be quantified by these relationships. Therefore, investigating the fundamental diagrams is an effective tool to assess pedestrian walking models, study pedestrian dynamics, and measure the practical performance of pedestrian facilities.

Flow is defined as the number of pedestrians passing a reference line in a unit of time and usually is computed for a unit of length (1 metre width). Hence, the unit of flow is *ped/m.s.* Density is defined as the number of pedestrians in one square metre of an area at a specific moment and is stated with ped/m^2 unit. Finally, speed refers to the average speed of pedestrians in an area at a specific time and its units are *m/s*. The fundamental traffic flow equation is expressed as follows:

$$q = k \times u \qquad \text{(Equation 2. 1)}$$

where q (*ped/m.s*) is the flow, k (ped/m^2) is the density, and u (*m/s*) is the mean speed. Several studies have been conducted to explore the fundamental relations (or diagrams); namely flow-density, speed-density, and speed-flow diagrams for pedestrian movements (*e.g.,* Fruin 1971;Seyfried *et al.* 2005; Helbing *et al.* 2007). All three diagrams provide the same information about pedestrian movements and each can be inferred from one another. Figure 2.1 shows the flow-density relationship taken from Daamen (2004).

Free speed (u^0) is the speed individuals tend to maintain while walking in uncongested and undisturbed situations (*i.e.,* low densities). The jam density (k_j) is the density of pedestrian crowd when the speed and flow are equal to zero. The maximum theoretical flow is known as capacity (q_c). Density and speed at capacity are denoted by Capacity density (k_c) and Capacity speed (u_c)

respectively. 'Stable region' is the region with approximately constant speed, which is very close to u^0. 'Unstable region' is the region where speed decreases with increase in density. 'Congestion region' is the part with the density higher than capacity density. The part with the density less than capacity density is 'free-flow region'.

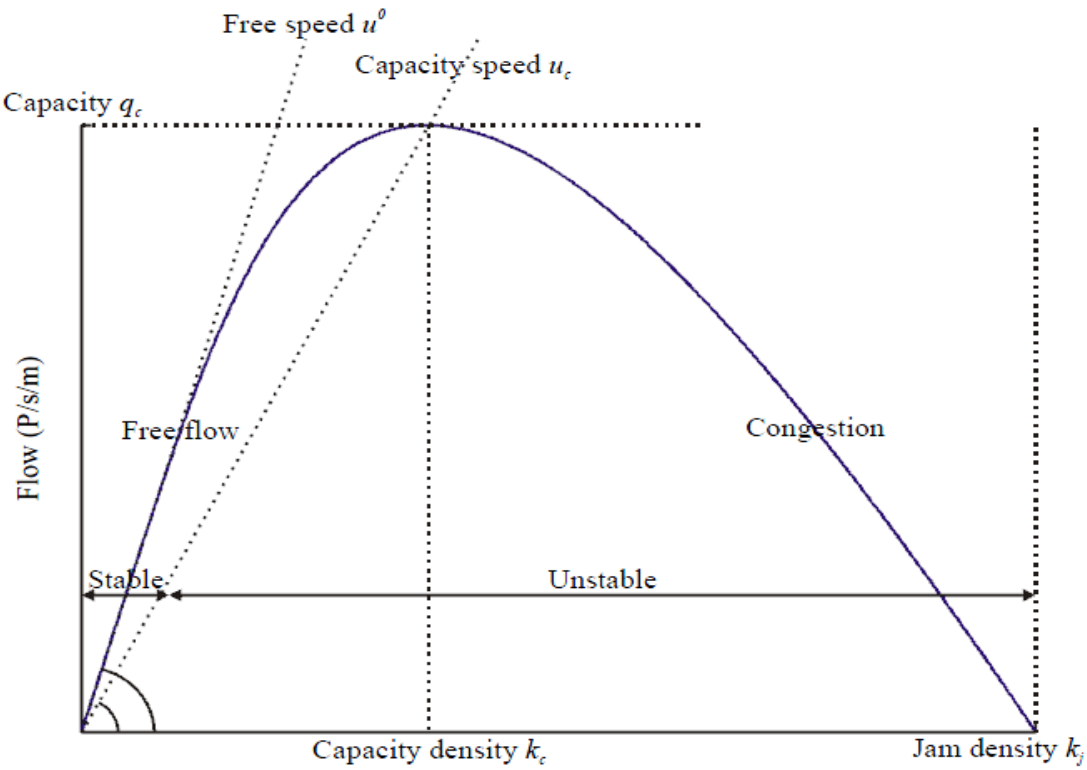


Figure 2.1: Assumed diagram of Flow-Density relation (source: Daamen 2004)

Pauls (1987) found that jammed pedestrian flows happen at density (jam density) of between 4 to 5 *ped/m*^{2} while Fruin (1971) reported the density of more than 5 *ped/m*^{2} as jam density. Weidmann (1993) estimated jam density of 5.4 *ped/ m*^{2} at which conflicts between pedestrians are not easily avoidable. The density of free-flow movement is found to be less than 0.5 *ped/m*^{2} (Pauls 1987; Fruin 1971). According to O'Flaherty & Parkinson (1972), the density at capacity (k_c) and capacity flow (q_c) are 1.89 *ped/m*^{2} and 1.29 *ped/m.s,* respectively. Pushkarev & Zupan (1975) reported 1.5 *ped/m*^{2} and 1.67 *ped/m.s* for capacity density and capacity flow, respectively.

When traffic demand is higher than available capacity, congestion occurs and that is when the flow begins to become over-saturated. According to Muramatsu *et al.* (1999) the jammed density is independent of the area size. Daamen (2004) stated that pedestrians' comfort quickly reduces at densities higher than 1 *ped/m*^{2}. Daly *et al.* (1991) found the capacity of 89 *ped/ m.min* for a corridor at a metro station in London. The higher value of capacity (92 *ped/ m.min*) was measured in the study of Lam & Cheung (2000) for a metro station corridor in Hong Kong. Lam & Cheung pointed out the smaller sizes of body and bigger tolerance for closer interpersonal distances among Asian people as the reasons for this.

2.1.2 Speed

Daamen (2004) reviewed observed speeds of pedestrians obtained from previous studies in uncongested corridors and found the average speed of 1.34 *m/s* with the standard deviation of 0.37 *m/s* for pedestrians in uncongested situations (free-flow speed). Speed of pedestrians approximately followed the Normal distribution and the mode speed of 1.2 m/s has been observed. According to Borghese *et al.* (1996) and Li *et al.* (1999) approximate speed of 2.2 *m/s* can be a conversion point from walking to running.

In most of the previous studies in the literature (*e.g.,* Fruin 1971) a linear relationship between speed and density is shown which indicates that as density increases individuals walk with lower speeds to avoid collision and prevent abrupt stops while walking. On the other hand, Weidmann (1993) fitted a S-bended curve to the data obtained from 25 different studies regarding this relationship. the free-flow walking situations happen in lower densities in the presence of non-moving (stationary) pedestrians or cross-directional flows in comparison to pedestrian flows with no crossing movements and no stationary individuals (Fruin 1971). Cross-directional flows are

those pedestrian streams which cross each other. The variation in speeds of individuals decreases when the density increases.

Speed of pedestrians can also be captured by studying pedestrian gait parameters. Step length is the distance from the point a foot touches the ground to the point that the other foot touches the ground and step frequency is the number of times that a foot touches the ground in one second. Stride length is the distance from the point a foot touches the ground to the point that the same foot touches the ground a second time. Stride frequency is the number of times the same foot touches the ground in one second. Therefore, stride frequency and length are half and twice of step frequency and step length, respectively. The following relationship was found between walking speed and step frequency and step length in the literature: Walking Speed = Step Frequency × Step Length. Therefore, it can be said that variation in the speed of a pedestrian is related to how she takes a step. Increasing speed can be performed by increasing step length and/or step frequency (Weidmann 1993). The results found in Crowe *et al.* (1996), Hediyeh (2012), and Hui *et al.* (2007) all show that walking speed is more related to step lengths rather than step frequency. Therefore, pedestrians prefer to increase (or decrease) their walking speed with increasing (or decreasing) their step length more than their step frequency.

In Hui *et al.*'s (2007) work, the mean of walking speed, step length, and step frequency for 1882 of Chinese individuals were found to be 1.22 *m/s*, 0.64 *m*, and 1.91 *Hz*. Živanović *et al.* (2005) measured step frequency for 1976 pedestrians and found the mean value of 1.87 *Hz* and the standard deviation of 0.186 *Hz*. Matsumoto *et al.* (1972) found the mean value of step frequency for 505 pedestrians to be 1.99*Hz* with standard deviation of 0.173 *Hz*. Finally, Venuti & Bruno (2009) reviewed previous studies and found the mean free walking speed to be within the range of 1.08 to 1.60 *m/s* , the mean step length of range 0.75 to 0.768 *m*, and mean step frequency of

1.82 to 2 *Hz*. Normal distribution was generally found for walking speed, step length, and step frequency in the literature (Hui *et al.* 2007; Hediyeh 2012; Matsumoto *et al.* 1972).

Fundamental traffic characteristics are influenced by individuals' characteristics and some external factors (Daamen 2004). According to Weidmann (1993) free-flow speeds can be influenced by three types of factors: pedestrian characteristics, movement conditions, and infrastructure characteristics. Age, gender, culture, body size, luggage availability, and route familiarity are some of pedestrian characteristics that can affect the speed of pedestrians. Many of these factors are correlated. According to Daamen (2004), what is influential in the study of walking behaviour are actually the factors related to age (*e.g.,* level of fitness or health condition) and not the age itself. However, age is used as a surrogate measure because observing and capturing this characteristic from pedestrian data is easier. Movement conditions contain among others trip purpose, presence of attractions, trip length, and weather conditions. And finally, the type of infrastructure (walkway, hallway, stairs, doors, and etc.) and surface conditions are among those characteristics of infrastructure that affect pedestrians' speeds (Lee 2005). These factors influence the speed of pedestrians and consequently affect the fundamental diagrams. However, discussing these factors in more detail is out of the scope of this thesis.

2.1.3 Fundamental Diagrams at Bottlenecks

Daamen (2004) investigated the flow-density relationship for pedestrian movement through a narrow bottleneck of width of 1 metre. Pedestrian Flow as well as the speed of pedestrians passing two reference lines (one inside the bottleneck and one upstream of the bottleneck) were measured for each time step. Then, the density at reference lines was calculated, using the traffic flow identity equation. It was found that the flow-density relationship for the total width

upstream of bottleneck cannot be described by a single fundamental diagram. Therefore, different fundamental diagrams for different widths have to be created. The reason is that speed and flow of pedestrians occupying different widths are not the same. For example while the flow is congested in the area which is laterally in front of the bottleneck, pedestrians in positions located further from (laterally) the bottleneck walk in almost free-flow conditions. These differences lead to the conclusion that different fundamental diagrams are required to explain macroscopic characteristics of the crowd upstream of the bottleneck in congested situations. Moreover, Hoogendoorn (2004) showed that for unidirectional flow, the capacity of a bottleneck does not have a linear relationship with its width. But instead, the capacity increases step-wise with bottleneck width and that is because of the so-called *zipper effect*, which allows lateral distance between pedestrians to be less than shoulder width.

2.1.4 Previous Studies on Differences between Fundamental Diagrams

Seyfried *et al.* (2008) compared some previous works and found jam densities (k_j) between 3.8 to 10 ped/m^2 and capacity densities (k_c) of 1.75 to 7 ped/m^2. Zhang *et al.* (2012) compared fundamental diagrams of two directional flows obtained from previous studies and found capacity densities (k_c) of 1.3 to 2.3 ped/m^2 and maximum flows (q_{max}) of 1 to 2 $ped/m.s$. These results show that there exist remarkable inconsistencies among different studies concerning pedestrian fundamental diagrams. Some of the reasons for these disagreements, proposed by researchers, are differences in cultures (Helbing *et al.* 2007) and pedestrians' different purpose of trip (Milinskii & Predtechenskii 1978). Figure 2.2 illustrates the fundamental diagrams of different studies taken from Zhang *et al.* (2012). The fundamental diagrams proposed by Weidmann are obtained from fitting the curve to 25 different studies.

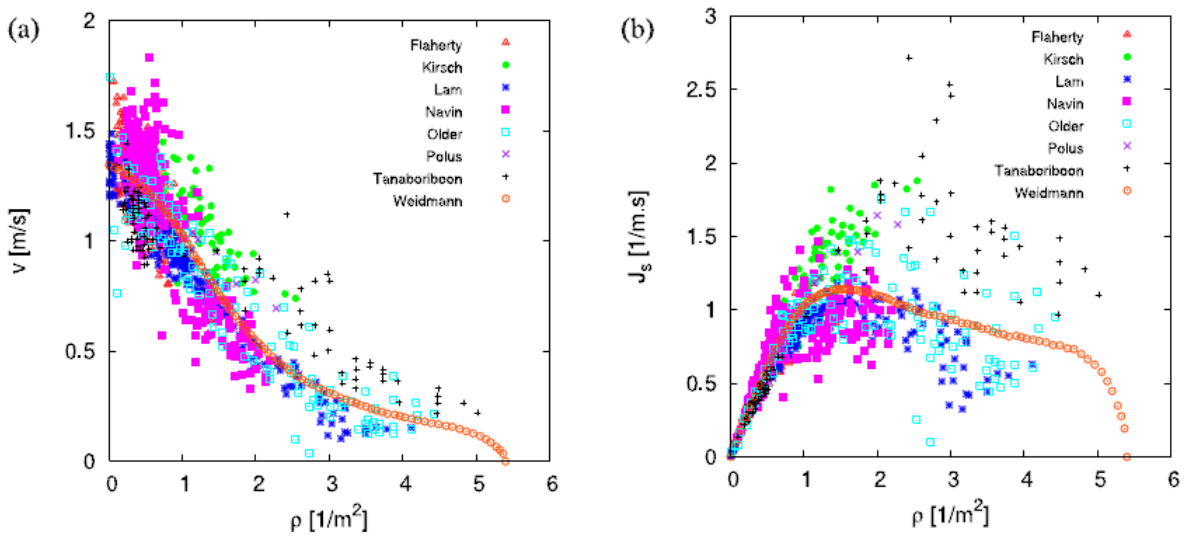


Figure 2.2: Fundamental Diagrams of two directional flows from different studies a) speed-density b) flow-density (Souce: Zhang *et al.* 2012)

Zhang *et al.* (2011) and Zhang *et al.* (2012) addressed the existing differences in fundamental diagrams in literature and investigated the effect of some potential reasons for these differences. Zhang *et al.* (2011) compared the fundamental diagrams obtained according to different measurement methods. Note that macroscopic pedestrian characteristic inside a system can be measured using many different methods. For example, flow and average speed passing a specific cross–section for each time interval can be measured and then the density over time is computed using traffic flow equation. Another approach is to measure the average speed over space and spatial density for a specific area and then compute the flow based on their relations. They found that fundamental diagrams resulting from different measurement methods have almost the same pattern and no obvious differences have been observed. They also found out that the fundamental diagrams for different facilities (*e.g.,* bottleneck vs. hallway) are not similar and therefore cannot be combined. Conversely, the fundamental diagrams for the same kind of facility but with different widths are found to be similar and therefore, can be depicted in one single diagram.

Seyfried *et al.* (2005) explored the fundamental diagram for a unidirectional pedestrian movement in a one dimensional system in normal situations. They obtained a linear relationship between velocity and the inverse of density and also discovered that for density values between 1 to 5 ped/m^2 the fundamental diagram matches well to that of pedestrian movement in a two dimensional plane. This implies that properties of two dimensional movement such as self-organized patterns and passing movements have no apparent effect on fundamental diagram.

According to Fruin (1971) there is no significant difference between fundamental diagrams of unidirectional and multidirectional flows. Conversely, Lam *et al.* (2002) found that effective flow decreases in response to the increase of imbalance in directional shares. Navin & Wheeler (1969) showed that a two directional flow reduces the capacity of the walkway by 4% when two opposite streams are evenly distributed (50%:50%). When the streams are proportioned at 10%:90 %, the capacity of the walkway will approximately be reduced by 14.5%. Zhang *et al.* (2012) compared fundamental diagrams of unidirectional and two directional flows obtained from previous studies and observed that for densities higher than 1 ped/m^2 the values of speed and flow are higher in unidirectional flows. However, as these diagrams were collected in various conditions and for different experiments it is hard to come to a conclusion about differences in unidirectional and bidirectional fundamental diagrams based on this comparison. They conducted experiments of two directional flow movements with different ordering degrees for densities below 2 ped/m^2 and found that the fundamental diagrams for all of them are very similar. They also discovered that the fundamental diagrams for unidirectional and bidirectional flows are clearly different; for densities higher than 1 ped/m^2 higher values of speed have been observed in unidirectional flows; yet, no clear difference has been found for densities below 1

ped/m^2. Maximum observed flow (q_{max}) was 2 and 1.5 *ped/m.s* for unidirectional and two directional flows respectively.

2.1.5 Self-Organized Patterns

In some specific circumstances, collective patterns of movement automatically form in pedestrian crowds. These patterns are self-organized in the sense that they are not planned in advance and they emerge in the crowd mostly for the purpose of avoiding collisions. Individuals tend to maintain a reasonable distance from each other and in particular from those walking in the opposite direction (Helbing & Molnar 1998; Ma *et al.* 2010). Helbing & Molnár 1995, Helbing *et al.* 1998, Blue & Adler 1999, Hoogendoorn & Daamen 2005, and Dzubiella *et al.* 2002, among others, have studied self-organization of crowds in the literature. In this sub-section, some self-organized collective behaviours are briefly described. Dynamic lane formation in counter flows, oscillation and arching at bottlenecks, and zipper effects are some of the well-known patterns to be explained.

When density exceeds a critical level in a crowd of bi-directional streams, pedestrians moving in the same direction form lanes and follow each other in order to avoid frontal collisions with opposite stream and walk more efficiently (Zhang *et al.* 2012; Helbing & Molnar 1998). The number of lanes is determined by the density of crowd as well as the width of the walkway (Helbing & Molnar 1998). As mentioned in the previous sub-section, the effect of bi-directional flow on the reduction of walkway capacity is not significant. One main reason for this is the potential formation of dynamic lanes in the presence of opposite movements. It was found that lanes form on the right side of the walkway with a moderate width (Daamen & Hoogendoorn 2003). When dealing with cross flows, the lanes of pedestrians form in the shape of strips or

groups of people walking in the same direction. Lane and strip formations help to reduce conflicts and the chance of pedestrians blocking each other and accordingly increase the average speed of pedestrians (Helbing & Johansson 2009).

Due to the *zipper effect* between two adjacent formed layers, the speed of these layers are approximately the same and are found to be about 1 *m/s* at capacity according to Daamen (2004). Layers are self-organized lanes that form in one directional flows. Zipper effect refers to the situation where the two layers next to each other overlap. Pedestrians in one layer occupy some space of the other layer with their shoulders as a result of body sway.

Daamen (2004) observed that when the flow through bottleneck is equal to capacity (when congestion happens) layers are formed and the observed lateral distance between centres of the layers is approximately 45 centimetres for both narrow and wide bottlenecks. The required lateral distance between pedestrians includes the width of a pedestrian's shoulder plus the distance for body sway. However, due to the zipper effect, the required distance is less than observed lateral distance between layers.

When in high density situations pedestrians with relatively high desired speed want to pass through a bottleneck (*e.g.,* a door), arching may occur. This phenomenon refers to the situation when pedestrians block each other due to the large number of conflicts between them. Consequently, nobody can pass the bottleneck and it becomes clogged and the crowd forms the shape of an arch in front of the exit. The extent of panic and the width of bottleneck are determining factors in the occurrence of this phenomenon (Daamen 2004).

Faster-is-slower pattern refers to clogging close to the exit. The average speed of pedestrians leaving the exit decreases when they attempt to walk faster. Consequently egress time increases (Saboia & Goldenstein 2012).

Oscillatory changes in movement direction are often observed in bi-directional flows at bottlenecks (*e.g.,* doors) at moderate densities. When a pedestrian in the crowd can pass through the bottleneck, other individuals with the same moving direction follow her. Therefore, the number and the pressure from standing pedestrians in this direction decrease and after a while the pressure of pushy pedestrians from other side (other movement direction) overcomes the current side. Also individuals from the current side may give way to the other side and consequently the stream in the current side stops and the pedestrians on the other side begin to pass through the door. This will result in a change in movement direction and the formation of oscillatory flows. This self-organized phenomenon reduces the delay and also decreases the number of conflicts between pedestrians with different directions (Helbing & Johansson 2009; Helbing *et al.* 2001).

In high density situations, *stop-and-go waves* and *crowd turbulence* are observed in crowds. When congestion arises and flow decreases to a very small value, the flow suddenly transforms to stop-and-go waves which may continue for some time. After this, the density increases to a higher value (*e.g.,* 7 *ped/* m^2 according to Fruin 2002) resulting in tremendously dense crowding. Consequently, stop-and-go waves abruptly transform to crowd turbulence where the random flows are generated and pedestrians are involuntarily moved with the crowd to any direction and are pushed by other individuals. As a result, some individuals may fall down and unless they stand on their feet very fast, they will be trampled (Johansson *et al.* 2008; Helbing *et al.* 2007).

2.2 Pedestrian Behavioural Levels

Pedestrian behaviour has been categorized into three different levels namely strategic, tactical, and operational levels (Hoogendoorn *et al.* 2002; Daamen 2004). An overview of these levels of behaviour is given in Figure 2.3.

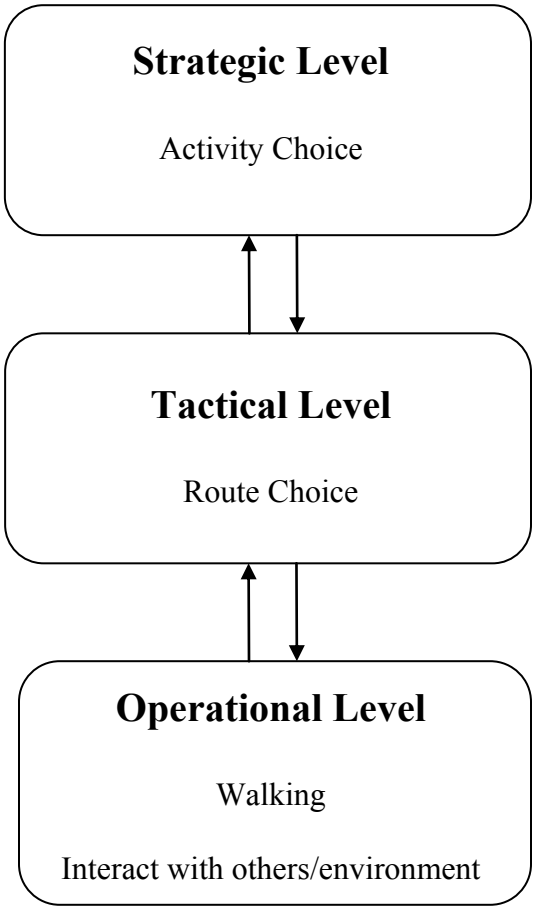


Figure 2.3: different levels of pedestrian behaviour

At the strategic level, long-term decisions are made. Pedestrians decide on the activities they intend to do. These activities can be either mandatory or optional such as buying a cup of coffee. Generally, decision process related to before-trip period takes place at this level. Short-term decisions are made at tactical level, taking into account the goals set at strategic level and based

on the information about the network and existing routes and conditions. These decisions include the performance order of activities selected at strategic level and activity scheduling, activity area choice, and route choice between the origin and the intermediate or final destination of pedestrian. The decision making process at this level is affected by two categories of external and internal factors. Internal or personal factors include pedestrian characteristics such as age, gender, attitude, trip purpose, and time-pressure. External factors include, among others, infrastructure and environmental conditions, presence of obstacles, timetables, as well as macroscopic characteristics of traffic flow such as average speed and congestion. At the operational level, the instantaneous decisions are made, in accordance with the objectives set at tactical level. Decisions at this level describe pedestrian walking behaviour including pedestrian's acceleration behaviour (*i.e.,* walk fast or slow), direction change behaviour, reactions to slow pedestrians or obstacles, deciding if wait or perform activities, and avoiding collisions. Interaction of pedestrians with each other plays an essential role at this level (Schadschneider *et al.* 2009; Daamen 2004; Hoogendoorn & Bovy 2001; Sahaleh *et al.* 2012) .

In this hierarchy of decision making, these levels are not completely segregated from each other. Decisions made at higher levels influence choices at lower levels. Besides, expected choices at lower levels affect the decisions to be taken at higher levels. For example, choices made at the strategic and the tactical levels constitute inputs for the operational level and influence the decisions made at this level. On the other hand, the expected traffic condition and consequently decisions at the tactical level are affected by collective walking behaviour at operational level; likewise, the decisions at the strategic level are affected by choices at the tactical level and vice versa (Ishaque & Noland 2008; Hoogendoorn & Bovy 2004).

It is usually assumed that strategic and tactical decisions are exogenous to the pedestrian simulation. Information from other areas such as sociology and psychology is needed to investigate the decision making process at these levels. However, in the literature exists only a few sources on activity choice set generation, activity scheduling, and activity location choice in urban areas (*e.g.,* Timmermans *et al.* 1992; Borgers & Timmermans. 1986). Nevertheless, a vast literature exists on route choice behaviour. Many studies have been focused on this important step of pedestrian behaviour *modelling*. Network and route characteristics such as the number of available routes, walking distance, route attractiveness, route straightness, crowd density , safety, surface condition, air and noise pollution, and the distance to obstacles, together with personal characteristics such as pedestrian's decision making style, age, and gender, as well as trip characteristics like the purpose of the trip are factors affecting route choice behaviour (Daamen 2004). Walking distance and walking time are the parameters being identified to greatly affect pedestrian route choice behaviour. Most route choice set generation approaches are on the basis of shortest or quickest path algorithms. Pedestrian route choice models are usually constituted of different types of discrete choice models and are based on random utility maximization theory. Borgers & Timmermans (1986); Hughes (2000); Hoffmann (2000); Hoogendoorn & Bovy (2004); Cheung & Lam (1998); and Daamen (2005), among others, have studied passenger route choice behaviour.

Strategic and tactical levels are considered exogenous to the model discussed in this thesis. Therefore, these levels of pedestrian behaviour are not discussed any further in the remainder of this thesis. The decisions made by pedestrian in order to arrive at the walking facility, as well as pedestrian arrival patterns, pedestrian route choice, and etc. have not been considered and are assumed known a priori.

The model described in this thesis addresses pedestrian behaviour at the operational level. Extensive research and different types of simulation models exist on pedestrian walking behaviour. Literature on models of this level of behaviour is presented in section 2.4.

2.3 Pedestrian Interactions

Pedestrians interact with each other and with the environment around them while walking. These interactions play an important role in the study of their walking behaviour. According to Goffman (1971) concept of s*canning,* pedestrians observe the environment when performing the task of walking. That is, each pedestrian is assumed to interact with those pedestrians in her scanning area and therefore will check on them and their movement. The scanning area is an ellipse which is larger in front of the pedestrian and smaller in the sides. However, the size of the scanning area changes depending on the traffic condition and density. Naturally, there is a strong cooperation between pedestrians that makes walking feasible (Wolff 1973). A single pedestrian faces several decision situations while walking. Even though these decisions are strongly related in reality, each of them can separately be associated to a behavioural mode (Bierlaire & Robin 2009). Three important types of interaction behaviours are explained in this section; Leader-Follower, Collision-Avoidance, and Group Behaviour are described in sub-sections 2.3.1, 2.3.2, and 2.3.3, respectively. Interaction behaviours may create some specific self-organized patterns of motion in human crowds such as formation of dynamic lanes.

2.3.1 Leader-Follower

Leader-follower pattern describes the tendency of an individual to adapt her speed and direction to another pedestrian to take advantage from the space the leader creates and to walk more easily and efficiently.

Reynolds (1999) proposed a leader-follower pattern which is the combination of *arrival* and *separation* behaviours. One or more individuals follow another moving individual assigned as the leader. Arrival behaviour describes follower's tendency to move towards a point located slightly behind a leader, called 'arrival target'. The follower adjusts her speed and direction while following the leader, that is, she slows down as the target gets near. In case she finds herself on the near future path of the leader (in a rectangular region in front of the leader), the follower move laterally away from leader's path and then again continues arrival behaviour. Also, followers themselves avoid crowding each other using separation behaviour.

Leader-follower can be modeled using the insight gained from the car following in driving behaviour models. Pipes (1953) first proposed car following behavioural model. The main concept is that the subject vehicle follows the leader and reacts to its actions. The leader is simply the vehicle in the front. Most car following models are built based on the sensitivity-stimulus framework, where a driver responses to stimuli from the environment. Normally, it is assumed that the leader's speed relative to the speed of the follower is the stimulus and the driver reacts to this stimulus by applying acceleration or deceleration. Many models have been generated based on different specifications of sensitivity term among which the linear car following model is the simplest one (*e.g.,* Herman *et al.* 1958; Chandler *et al.* 1958) Gazis *et al.* 1959 proposed a nonlinear car following model in which the reaction to relative leader speed is inversely related to the spacing between the leader and the subject vehicle. This model can be regarded as the most well-known model in this context. Various other car following models have been suggested in this subject matter. For example, Komentani & Sasaki (1958) proposed a model capturing the tendency of the follower to maintain a minimum safe distance with the leader to avoid collision. Helly (1961) suggested a car following model based on the driver's

propensity to minimize the difference between the actual and a desired space headway with the leader in addition to minimizing the leader relative speed. Newell (1961), Bando *et al.* (1995), and Aycin & Benekohal (1998) are some of other examples.

According to the choice-based walking behaviour suggested by Bierlaire & Robin (2009) , two types of choices can be modeled regarding leader-follower behaviour. The first one is the decision of choosing a leader (or not to follow a leader). The second one is the decision on how to react to the leader's behaviour if so is applicable. Characteristics of the crowd around the decision-maker (such as density and speed) and also the potential leaders' behaviours can influence the choice of a leader. Those individuals in decision-maker's visual field with the behaviour similar to that of the decision-maker, especially with regard to direction and speed, are more likely to be chosen. Regarding the second type of choice, Robin *et al.* (2009) proposed that pedestrian's speed and direction choice is influenced by the leader. They modeled the leader-follower behaviour based on the sensitivity-stimulus framework which was derived from the car following models in the literature and has been adapted to pedestrian movement case. Leader's behaviour can also affect other choices like route choice (Bierlaire & Robin 2009).

Leader-follower pattern has also been used in *modelling* group behaviour. For instance, Li *et al.* (2001) proposed a leader-follower framework for *modelling* groups of crowd in which each group is guided by a leader. Similarly, in Loscos *et al.* (2003) work, decisions about the movement of the group are taken by the leader and group members follow the leader. Ji & Gao (2006) proposed crowd evacuation behaviour based on a *dynamic* grouping process. The crowd is consisted of different groups of people where each group has a leader and the leaders have to find the exit location for their followers. In this model, followers follow the nearest leader. Therefore, the grouping process is dynamic and each follower may be guided by different leaders

during evacuation. The authors also examined the effect of having more than one leader on the efficiency of egress process. They found that total egress time decreases when more leaders are included, but only up to a saturation point. Beyond this point, the efficiency of evacuation process may decrease with increasing the number of leaders.

2.3.2 Collision Avoidance

Contrary to leader-follower behaviour, in collision avoidance behaviour, a pedestrian is negatively affected and repulsed by another individual or object (Bierlaire & Robin 2009).

Personal space is one of the main concepts in social interactions and psychology which is also relevant in constrained behaviours of pedestrians; especially in collision avoidance context. Each pedestrian maintains a certain distance from other individuals and also from obstacles and boundaries such as walls and columns. *Personal space* can be defined as the area surrounding an individual's body which is considered to be exclusive space of the individual (is often described by subjective ownership). Invasion of this space by other individuals is considered to be a violation of self boundaries and can make the person feel discomfort, anger, or anxiety (Hall 1966; Dosey & Meisels 1969; Sommer 1969). It also has been suggested that *personal space* serves as a protective mechanism which is built based on the ability of the individual to perceive signals from her social and physical environment and expands in response to threat. Its role is to generate the spacing pattern that governs distance between people and on which individual behaviours are relied (Dosey & Meisels 1969; Webb & Weber 2003; Horowitz *et al.* 1964). Hence, *personal space* is a buffer zone to protect the person as well as a psychological part of the person herself.

The effects of personality characteristics and social situations on the physical distance a person wants to maintain with others have been confirmed in several studies in psychology. For example, variables such as sex (Hartnett *et al.* 1970), and cultural background (Hall 1966) have been found to impact personal space. Furthermore, it has been found that people experiencing tension will create a bigger interpersonal distance with other individuals (Dosey & Meisels 1969). Frankel & Barrett (1971) showed that people with low self-esteem require a greater interpersonal distance compared to those with high self-esteem and also that people establish different personal spaces facing with individuals from different races. Newman & Pollack (1973) found that aggressive and defiant adolescents create greater personal space than normal and non-disruptive adolescents. Kuethe (1962) found that with increasing age people require greater distances. Bailey *et al.* (1972) pointed out that a male person generates a greater interpersonal distance with other individuals than a female. Therefore, it is evident that individual characteristics including, among others, sex, height, age, and body size may affect individual's perception of space, interpersonal distance, human interactions, and consequently walking behaviour. However, capturing these characteristics from real world scenarios is not an easy task to do and sometimes requires an analyst to conduct a controlled experiment.

Pedestrians are more affected by intrusion to their personal space from the front. The size and the shape of the personal space are continuously changing based on the crowd density and the travel speed of the pedestrian (Karamouzas *et al.* 2009) and it is also dependent on individual's head orientation and visual system (Hayduk 1983). Goffman's personal space is an oval which is narrow to the sides of the person and long in front of her (Goffman 1971).

Fruin (1971) recommended the plan view of human body as an 18 by 24 inch (2.3 square feet) ellipse (approximately 50 by 60 cm) for standing areas which represents the body depth and

shoulder breadth measurements as shown in Figure 2.4. This area defines the actual body space for the average individual.

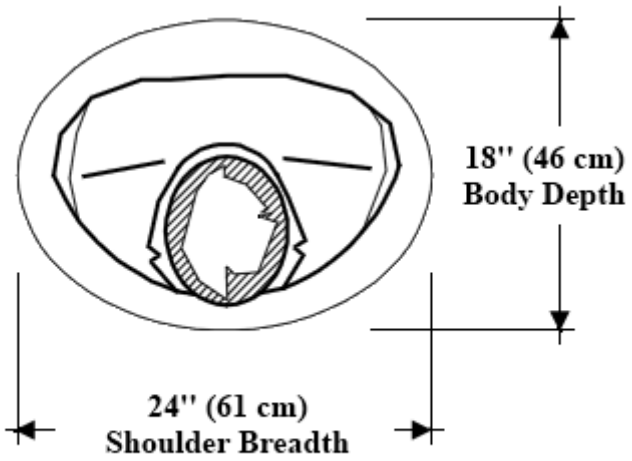


Figure 2.4: Fruin's body ellipse for standing areas (Source: Fruin 1971)

People tend to avoid conflicts with others unless such crowding is not avoidable (Fruin 1971). Adequate space is required both laterally and longitudinally to avoid physical contacts. The required lateral spacing is defined by the width of the human body (shoulder) together with allowance for body sway and longitudinal spacing is identified by pacing distance plus perception and reaction times. It should be noted that vision is an important factor in perception of space and avoiding collision. Based on the aforementioned factors, Fruin (1971) has considered a lateral space of 28 to 30 inch (71 to 76 cm) and a longitudinal spacing of 8 to 10 ft (2.5 to 3 m) resulting in a minimum personal area of 20 to 30 ft^2/pr (2 to 3 m^2/pr) to move comfortably and to avoid contacts with others in unhindered walking situations. According to Wolff (1973), interacting pedestrians (especially those of the same sex) not always try hard to avoid the contact with each other. But instead, they carry out *step-and-slide* movement. That is, they turn their shoulders and walk aside just slightly. However, the chance of touching between

pedestrians is still present in this movement. The personal space of a walking pedestrian is different from personal space in stationary situation. Pedestrians require larger distance in front while walking to avoid collisions. The spacing required in front and the lateral distance needed are, respectively, smaller and larger in stationary situation (Bandini *et al.* 2012).

However, interpersonal distance decreases as the density increases. Much smaller personal areas are perceived in crowded situations where movement is restricted, resulting in less freedom of motion. Collision avoidance behaviour only takes place when pedestrians are sufficiently close to each other in dense crowds. On the other hand, when the density decreases or average relative speed increases people perform collision avoidance behaviour earlier (Osaragi 2004).

Excessive level of crowding results in inevitable contacts with other pedestrians and can disturb pedestrian comfort and in more severe case result in mass fatalities. Fruin (1971) suggested a minimum desirable occupancy of 5 to 10 ft^2/pr (0.5 to 1 m^2) where physical contact with others is avoidable. According to Bandini *et al.* (2012) 40x40 cm is the space individuals occupy in high density conditions. Jelić *et al.* (2012) suggest that in high density situations people walk in lockstep and the distance they keep from each other turns into the distance they move forward when they take a step.

Several methods have been proposed to analyze and model collision avoidance behaviour among nearby pedestrians including among others: *force-based methods* (Helbing & Molnár 1995; Sud *et al.* 2007; Heïgeas *et al.* 2010) and *Geometric-based approaches* (Paris *et al.* 2007; Van Den Berg *et al.* 2008; Fiorini & Shiller 1998; Feurtey 2000; Karamouzas *et al.* 2009; Guy *et al.* 2009). Helbing & Molnár (1995) first proposed social force model based on the laws of Newtonian mechanics where each pedestrian corresponds to a particle. Since then, many social

force models have been suggested to simulate pedestrian crowds in normal and emergency situations (*e.g.,* Pelechano *et al.* 2007 and Helbing *et al.* 2005). Social force models simulate pedestrian walking behaviour as interaction between particles. The main concept of all these models is that pedestrians employ repulsive or attractive forces on each other. The repulsive social force captures the tendency of individuals to keep a minimum distance from other individuals. Pedestrians just interact when they get close. Therefore, the resulting movements tend to become unrealistic and oscillations may occur. In crowded and cluttered situations, this may cause pedestrians to get stuck behind obstacles or may cause them to change their orientations constantly, push each other, and move back and forth as there is no beforehand anticipation in these models and only information at local level is considered. Alternatively, many other collision avoidance approaches have been proposed based on collision prediction in which pedestrians' trajectories are extrapolated to anticipate and avoid collisions in the near future. For example, Feurtey (2000) proposed a space-time algorithm to predict potential collisions with other agents and modify pedestrians' trajectories and/or speed in (x, y, t) space. Paris *et al.* (2007) proposed a geometric based collision avoidance model in which agents perceive the trajectories of nearby pedestrians to predict potential collisions. Likewise, Shao & Terzopoulos (2007) presented a set of reactive behaviour regimes to lead pedestrians in a collision-free path. Furthermore, moving multiple agents at the same time can be modeled using *prioritization rules* which is based on successive navigation planning in a way that the pedestrians whose path have been planned earlier become potential moving obstacles for the following ones (*e.g.,* Lau & Kuffner 2005; Foudil & Noureddine 2006). However, this method does not always result in optimal navigation paths. *Velocity Obstacle* concept has also been employed to model the reaction of agents to dynamic obstacles (*i.e.,* other agents) (*e.g.,* Fiorini &

Shiller 1998). Based on this technique, an agent (a_i) avoids collision with other agents' (a_j) trajectories (or their extrapolated trajectories) by selecting a velocity outside the velocity obstacle. Velocity obstacle consists of all velocities of agent a_i which result in collision with agent a_j walking at velocity vj. This method may lead to oscillation in crowded environments as it assumes that other pedestrians (a_j) are just moving obstacles and discards the fact that they are decision-makers themselves as well. *Reciprocal Velocity Obstacle* has been proposed to address this problem in which each agent is a decision-maker and reacts to the presence and movement of other agents (Van Den Berg & Manocha 2008). Another collision behaviour method is suggested by Bierlaire & Robin (2009). According to the choice based walking behaviour proposed and similar to their leader-follower pattern, the choice of a collider and how to react to the collider's behaviour are two types of choices that have to be modeled. The process of determining a collider is similar to that explained for leader: the characteristics of the surrounding crowd and the behaviour of the potential colliders affect the choice of a collider. Individuals in the decision-maker's visual field with a direction and speed close to those suggested for a potential collision are more likely to be chosen. Robin *et al.* (2009) suggested that a pedestrian changes her direction in response to the existing collider. Interactions in this model are not limited to local level. Rather, pedestrian's visual fields are taken into account in defining collision avoidance behaviour.

While walking, individuals come across different scene elements such as cars and stationary obstacles like curb lines, doors, and walls. Pedestrians naturally keep a certain distance not only with other pedestrians but around all objects and try not to collide with them. Himanen & Kulmala (1988) modeled interactions between pedestrians and drivers using a discrete choice framework in which pedestrians can choose to stop or pass. Vehicle speed and vehicle size are

among the variables that affect the choice in this model. When pedestrians cross the road they also may interact with dangerous elements like streetcar lines. Pedestrian can choose to pass or stop in these situations. The choice is influenced by pedestrian's characteristics and level of danger, which for example can be measured in terms of vehicle speed (Bierlaire & Robin 2009). Corners can be other scene elements in the way of pedestrians. Pedestrians keep *shy-away* distance from walls, buildings, and curbs. This distance is dependent on the wall condition (*e.g.,* wall materials) as well as the movement condition (*e.g.,* bi-directional flows, obstacles, density) (Daamen 2004). Pedestrians may walk away from the wall to enhance their vision and decelerate (or stop) to check the possibility of a collision at the crossing. Likewise, pedestrians at the door can stop, decelerate, or change their trajectories to anticipate and avoid collisions. Moreover, Pedestrians may go around by left or right, turn back, or stop when come across stationary obstacles or static pedestrians. Pedestrian's characteristics and traffic density are two important factors, among others, affecting decision making process in all aforementioned scenarios.

2.3.3 Group Behaviour

In social science, a group is described by two or more humans who interact with each other, have relatively homogeneous characteristics, and collectively share a sense of unity (Forsyth 2010). Expectations and obligations for group members within the group is what distinguish them from aggregates. Group formation has important effect on pedestrians' behaviour. A pedestrian crowd consists of people walking individually and also those walking in groups (Aveni 1977). According to Aveni (1977) and Coleman & James (1961) the larger portion of pedestrians in a crowd walk in groups rather than walking alone. Most of people at public events, commercial streets, shopping malls, and museums form groups with their friends or family members. Pedestrians in the same group maintain the group while walking. They stay together and keep a

short distance from each other while pedestrians belonging to different groups keep away from each other. Groups made up of a large number of pedestrians may hinder and obstruct other pedestrians and consequently influence the characteristics of crowd flow such as speed or density. Therefore, the study of group behaviour is of great importance in providing a realistic walking behaviour model.

While there has been a great deal of work done on simulating crowd behaviour, only little work has studied the impact of grouping phenomenon on crowd dynamics. A brief introduction to some related work is discussed below.

Reynolds (1987) created a model for simulating bird flocking using three steering behaviours: separation, cohesion, and alignment. It is assumed that birds' attempts to maintain the distance to the flock (group) centre, at the same time to avoid collision with other group members, as well as matching the velocity with other birds in the flock leads to synchrony of flocking behaviour. Yang *et al.* (2005) have proposed a two-dimensional cellular automata model to simulate and investigate the effects of kin behaviour on evacuation process. They found that evacuation efficiency is highly affected by the number of sub-groups and people in each sub-group. That is, if there is a great number of sub-groups or a large number of people in each sub-group, the evacuation effectiveness will highly be reduced. Klüpfel *et al.* (2005) investigated how walking speed changes with the change in group size and they observed that as the group becomes bigger, walking speed decreases. Sarmady *et al.* (2009) developed a CA-based model using the principle of least effort (PLE) to simulate the movement of pedestrians that are part of a group. This model is built based on the idea that the group members tend to maintain short distances with other members or the leader of the group. The least effort principal suggests that people try to reach their goals by making the least amount of efforts (Zipf 1949). Regarding human crowd motion,

this implies that people walk through the environment by minimizing the effort in terms of time, distance, congestion, or the change in speed. Qiu & Hu (2010) used agent-based crowd simulation to model different group structures. The intra-group structures (*i.e.,* relationship among the members inside a group) and inter-group relationships (*i.e.,* relationships among different groups) are captured by influence matrices. In this model, group behaviour is maintained using two aspects of movements: aggregation and following. Aggregation is when a group member moves toward the centre of other members in the group who have influence on her and following happens when an individual follows an average moving direction of other group members who have influence on her. It was found that crowd behaviour is highly influenced by different group sizes, intra-group and inter-group relationships. However, the model was not validated on real data. Moussaïd *et al.* (2010) extended an existing social force pedestrian behaviour model to investigate the effects of social interactions between group members on crowd dynamics. This individual-based model shows that social groups greatly impact human crowd behaviour and traffic efficiency. It turns out that at low densities group members tend to walk side-by-side. But at higher densities, the walking pattern bends forward into V-shape or U-shape patterns due to less available space in these cases. These spatial patterns may have resulted from each individual's tendency to communicate with other members inside a group. Qiu & Hu (2010) developed an agent-based model of pedestrian dynamic groups based on both utility theory and social comparison theory. Utility function is used to simulate group behaviours of pedestrians and social comparison theory is used to study interactions between agents.

Socio-psychological theories such as five-factor personality theory or the social comparison theory are widely employed to simulate groups in human crowds in recent studies (*e.g.,*

Durupinar *et al.* 2008; Ghasemaghaee & Oren 2007; Jaganathan *et al.* 2007; Fridman 2007). The simulation of social groups has also been of attention in animation industry. In this area, a group is defined by a number of people in the same physical environment who share common goals. Musse & Thalmann (1997) have proposed a model of crowd behaviour inside virtual environment relied on group inter-relationships and formed by autonomous agents. In this model, pedestrians belong to a set of goal-directed groups where members of the same group share the same list of goals and each group has a leader. Each pedestrian is identified by an emotional status, a level of dominance, and a level of relationship with other groups. Pedestrians react based on their own emotional parameters and consequently they may leave their groups and join other groups. Villamil *et al.* (2003) presented a model to generate groups based on interactions among agents and similarity of ideas between members of a group. Each pedestrian is characterized with a set of parameters, some of which change during each interaction. Also, each agent is associated with some abilities, including movement, perception, interaction, and memory. Each Pedestrian selects those agents having good quality of interactions in order to create a group with them.

Highway Capacity Manual (HCM 2010) does not contain any material on groups of pedestrian who are socially bond. It only provides some information about the level of service (LOS) in platoons. Platoons are defined as groups of pedestrians who walk together unintentionally; they form in pedestrian crowds as a result of external factors such as signal control or when the passageway is impeded because there is not enough space and consequently, faster pedestrians get stuck behind slower pedestrians. HCM reports that usually LOS in platoons is one level lower than the average flow criteria. According to HCM, for walkways and sideways, the

jammed flow (LOS F) starts at 1 m^2/ped in platoons. However, platoons are not formed with social bonds and therefore are not of interest in this research.

2.4 Pedestrian Walking behaviour Models

Modelling pedestrian behaviour is relatively a difficult task. Driver behaviour in vehicle-based transportation systems, where vehicle flows are normally limited to one-dimensional flows, is suggested by pre-defined corridors of the road environment and is controlled with specified traffic rules. In contrast to driver behaviour, walking is a complex behaviour which cannot simply be described by geometric representation of transport facilities and lanes. Human movements are two-dimensional and tend to be affected by a large number of parameters. Walking behaviour is described to be highly unconscious and difficult to predict. It may greatly be affected by human factors which vary widely across individuals. It can be argued that each pedestrian has a unique walking behaviour. Freedom of movement in public places and having the chance to carry out discretionary or recreational activities while walking provides pedestrians with a large number of alternatives to choose from. Furthermore, a complex process of decision making including the perception of environment by pedestrian and processing and evaluating the information is required in order to make a decision (Papadimitriou *et al.* 2009; Camillen *et al.* 2009; Hoogendoorn 2003).

A wide range of approaches have been proposed to model pedestrian behaviour at different behavioural levels in the literature. The focus in this research is on modelling pedestrian behaviour at operational level. The goal is to select a proper model which can provide a realistic representation of pedestrians' behaviours. Some of the examples of the proposed models are listed here. For example, Perez *et al.* (2002), Varas *et al.* (2007), and Blue & Adler (2001) used

cellular automata to model pedestrian behaviours. Tajima & Nagatani (2001) and Muramatsu *et al.* (1999) proposed navigation models based on lattice-gas models. Helbing & Molnár (1995) proposed a social force model. Hoogendoorn *et al.* (2002) modeled the path choosing of pedestrians by optimizing a specific cost function.

It should be noted that pedestrian's physical characteristics (especially age and gender), travel purpose, and type of infrastructure are identified as factors affecting pedestrian walking behaviour. However, they are rarely taken into account in walking behaviour models. Also due to the cultural differences that exist among societies, walking behaviour varies among different countries and individuals do not walk in the same way.

Pedestrian modelling approaches have been categorized based on different criteria in the literature. For example, classification into rule-based versus force-based models, microscopic versus macroscopic models, or continuous-space versus discrete-space models (Schadschneider *et al.* 2009; Schadschneider and Seyfried 2009). However, the most commonly used category is macroscopic versus microscopic models.

In microscopic models each pedestrian is an agent and her walking behaviour is studied separately, taking into account the interactions between agents and other pedestrians in addition to the environment around them. However, sometimes complicated programming process and unexpected errors involved with these approaches will bring complications to the model which directly affects the efficiency of it (Shiwakoti & Nakatsuji 2005). *Modelling* pedestrians at microscopic level requires an accurate representation of space and individuals' characteristics (Klüpfel 2012). This category of models can itself be subcategorized to continuous-space and discrete-space models.

Continuous-space models are capable of modelling continuous movements of pedestrians. In these models continuous status are assigned to space variables (these variables are represented by real numbers). Conversely, space is characterized by discrete variables in discrete-space models. The method of model execution determines whether to employ a continuous or discrete space model (Klüpfel 2012; Chraibi *et al.* 2009). Cellular Automata models are the best known discrete-space models (*e.g.,* Zhao *et al.* 2008; Varas *et al.* 2007; Blue & Adler 2001; Burstedde *et al.* 2001; Fukui & Ishibashi 1999). Discrete choice walking behaviour model proposed by Robin *et al.* (2009) is a discrete-space model in which the pedestrian selects her position at next time step among a set of discrete positions. This model is described in detail in chapter 3. Social-force models are widely used continuous models (*e.g.,* Helbing & Molnár 1995; Helbing *et al.* 2000).

Macroscopic models do not capture individual behaviours. These approaches model the large crowd as a whole based on the resemblance of aggregate crowd movement to continuous fluid or gas. The system is described by aggregate characteristics such as flow, average speed or density. Disregarding individual pedestrian behaviours in these approaches leads to not capturing some key issues in the system such as congestion occurrence or the possible low level of pedestrian safety in the crowd (Shiwakoti & Nakatsuji 2005). Moreover, interactions between individuals cannot directly be investigated based on these models. Macroscopic models are generally in the category of continuous-space models. Helbing (1992), Henderson (1974), and Hughes (2003) have used macroscopic approaches to model pedestrian behaviour among the others.

A quick review of some practical macroscopic models is given in sub-section 2.4.1. Then, in sub-section 2.4.2 two well-known microscopic models, cellular automata (CA) and social force

model (SF) are briefly reviewed. Another microscopic approach to model pedestrian behaviour using discrete choice modelling (DCM) framework is described in detail in the next chapter.

2.4.1 Macroscopic Models

In macroscopic models, density and the average speed of the crowd are described using partial differential equations derived from fluid or gas dynamics. The macroscopic models are not the topic of this thesis. In this sub-section only some general features of the fluid dynamic based model proposed by Helbing (1992) is briefly explained. In this model each pedestrian has a preferred direction of movement and pedestrians with the same intended direction of movement generate a type of motion. Individuals of each type of motion are specified by their position, current velocity, and their intended velocity at each time. The number of pedestrians in each type of motion in a specific area defines the density of that type of motion. Differential equations describing average speed, variance of speed, and spatial density are extracted from these calculations. Interaction between individuals, pedestrians' desire to reach their intended velocity, increase and decrease in density at entrances and exits, and also presence of those pedestrians that change their type of motion in order to turn to left or right are the factors that justify the changes in the densities of motion types in the model.

2.4.2 Microscopic Models

More attention has been focused on microscopic models over the last years. This is mainly because of the improvement in computational techniques which allows a researcher to develop a more detailed and sophisticated model. Compared to macroscopic models, these models provide an in-depth and detailed representation of behaviours by capturing interactions between pedestrians. In this sub-section a brief review of these approaches are given.

Cellular Automata

Cellular automata models first were used in vehicle traffic simulation context. And recently they are used to model pedestrians walking behaviour as well. In order to generalize CA model to pedestrian context, first it should be adapted to pedestrian bi-directional movement rather than unidirectional movement in vehicle traffic case. Blue & Adler (1998) proposed a bi-directional CA model and then extended it to four directional movements to capture crossing transitions as well (Blue and Adler 2000). Time, state variable, and space are discrete variables in cellular automata models. The space is modeled as regular grid cells and pedestrians are entities (particles) that occupy these cells. The cells are usually squares of 40 centimetres long (the area required by an individual in high density situations). The time is discrete in a sense that the coordinates of agents are updated every time step. The time interval is normally specified by the reaction time of individuals. The state of each cell is updated every time step, which is one or zero if the cell is occupied and not occupied respectively. The cell occupancy is identified based on specific rules applied to cell's previous state and the previous state of its immediate neighbourhood (neighbour cells). That is, the probability that each entity (pedestrian) hops to each of the cells in its neighbourhood is dependent on the position of other individuals in its surrounding area. Therefore, this model is governed by localized rules. The movement of each Pedestrian consists of lane changing and cell hopping. Figure 2.5 illustrates the possible directions of movement and the probability of hopping to each of the cells for a Von Neumann neighbourhood case.

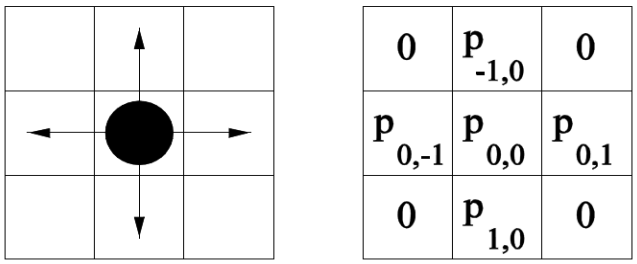


Figure 2.5: the directions of movement and probabilities for a Von Neumann neighbourhood case (source: Schadschneider and Seyfried 2009)

Two parallel stages are employed in each time step: lane changing (side stepping) and stepping forward (assigning velocity). Some sets of rules are used for this purpose such as gap identifying or eliminating conflicts by addressing some special cases such as two pedestrians that are laterally adjacent and may or may not side step on each other. Gap is the number of unoccupied cells ahead. First, if any of the left or right adjacent cells of the pedestrian is empty, the pedestrian is assigned (among the current cell and adjacent cells) to the lane with the maximum gap. Then, the forward movement is made by allocating a speed to the pedestrian depending on the available gap. Different CA models vary according to the specified rules. *Floor field model* is a special case and a more sophisticated type of CA models. It captures larger ranges of interactions and effects rather than just local effects in particle's neighbourhood area and therefore is more realistic in reproducing actual behaviour of pedestrians. However, the description of this model is out of the scope of this study. A more thorough description of this model can be found in Kirchner *et al.* (2002) and Burstedde *et al.* (2001) . Burstedde *et al.* (2001) found that Self-organized patterns such as lane formation can successfully be simulated by employing this type of models without explicitly modelling interactions between pedestrians.

CA approaches try to define dynamics of pedestrians by simple rules which reflect psychological factors of pedestrian movement. However, the rules identified do not completely demonstrate

and reflect the behaviour of pedestrians in actual situations. Moreover, the discrete space model brings about some limitation in the movement of pedestrians (pedestrians actually jump from one cell to another cell). The static (fixed) discretization of the grid cells is also a limitation in this approach. Sahaleh *et al.* (2012) stated that CA models have never been validated at microscopic level.

Social Force Model

Social force models are the most widely-used pedestrian models. They are implemented in popular simulation tools such as VISSIM. Social force models are force-based models that model pedestrians as particles which interact with each other. The interactions of pedestrian with other individuals and also with the environment are modeled with attractive and repulsive forces using Newtonian equations. Time and space are continuous variables. Each Pedestrian is represented as a particle with radius r_i that wants to reach a specific destination. The effects of other individuals and geometry on a way to the destination are captured in this model. The equation of motion is as follows:

$$m_i \frac{\overrightarrow{dv_i}}{dt} = \vec{f_i}\,(t) + \vec{\varepsilon_i}(t) \qquad \text{(Equation 2. 2)}$$

where $\vec{f_i}\,(t)$ is the summation of forces affecting pedestrian i, $\vec{\varepsilon_i}$ is the fluctuation which captures the random effects on individual's behaviour, m_i is the mass of the pedestrian, and $\overrightarrow{dv_i}$ is the velocity of pedestrian i.

The forces applying on pedestrian i is defined as follows:

$$\vec{f}_i(t) = \frac{m_i}{\tau_i}\left(v_i^0(t)\vec{e_i}(t) - \vec{v_i}(t)\right) + \sum_{i\neq j}[\vec{f}_{ij}^{soc}(t) + \vec{f}_{ij}^{phys}(t) + \vec{f}_{ij}^{att}(t)] + \sum_b \vec{f}_{ib}(t) + \sum_k \vec{f}_{ik}^{att}(t) \quad \text{(Equation 2. 3)}$$

The first term is the driving force describing the tendency of pedestrian to move towards a given destination by adapting her speed and direction with the desired velocity during a specific relaxation time τ. v_i^0 is the desired speed and $\vec{e_i}$ is the desired direction of pedestrian. Relaxation time is similar to reaction time (time step) in discrete models. The first summation term captures the interaction between the individual and other pedestrians. It includes repulsive social forces, physical forces, and attractive forces between individuals. The social force is the repulsive force from pedestrian *j* on pedestrian *i* and captures the tendency of pedestrians to maintain a certain distance from each other. The repulsive physical force prevents pedestrians from overlapping (high density situations impose collision of pedestrians). It consists of a body force and a sliding friction force (for more details refer to Helbing *et al.* 2000). The second summation is the repulsive force from borders and describes the tendency of pedestrians to keep distance from obstacles such as walls. This term also contains the physical force which keeps individuals from overlapping with obstacles. Both the repulsive social force and repulsive force from boundaries decay exponentially as a function of the distance. Finally, the last term captures the fact that pedestrians may be attracted by scene elements such as shops, advertisements, and etc. Then, the second-order partial differential equations resulted from Equation 2.3 should be numerically solved.

One concern about this model is that most previous validation processes performed on SF models are qualitative rather than quantitative validation based on real empirical data (Chraibi *et al.* 2009; Teknomo 2002). The SF model has been validated qualitatively at macroscopic level

based on self-organized patterns such as oscillation in bottlenecks, faster-is-slower effect, and clogging in front of exits. However, it has not been validated quantitatively at microscopic level according to individuals' walking behaviour (Sahaleh *et al.* 2012).

Saboia & Goldenstein (2012) address some problems in social force model. It has been argued that in some conditions, such as low density situations, pedestrians do not act rationally but rather act like irrational particles. For example, they move repeatedly towards the obstacles rather than finding their way around the obstacle. That is because the desired direction of movement is towards the destination even though there is an obstacle between pedestrian's current position and the destination. The authors addressed this problem by changing the driving force component in SF model and allowing pedestrians to change their desired direction according to their perception from environment by using mobile grids. It was shown that the new model results in more coherent trajectories compared to the original SF model and is also able to generate self-organized patterns like the original SF. Another problem with this model is that the social and physical repulsive forces do not assure that pedestrians do not collide with each other and in some situations they may overlap or share the same space (Teknomo 2002; Lakoba 2005). Moreover, in passing through bottlenecks, if individuals near the exit have a few seconds of delay in passing the door, pedestrians at the farthest distance from the exit will turn around and move away from the exit (Lakoba 2005).

Dealing with huge and complicated geometrics such as shopping malls is more difficult in social force models compared to CA models (Klüpfel 2012). Social force models are more sophisticated in terms of computational complexity compared to cellular automata models. But on the other hand, they are more flexible and reasonable in terms of space and time because of the continuous specification of these variables.

Compared to behavioural approaches which focus and make use of individual characteristics and preferences to establish mathematical models that describe pedestrians' behaviour, social force model is more of a physics-based model rather than focusing on behaviours of individuals. Forces in SF models are specified based on physical laws and then they have been employed to explain the behaviour of pedestrians. There are also microscopic models that focus on behavioural aspects of pedestrian movement using discrete choice modelling which will be explained in the next chapter.

2.5 Summary

Pedestrian behaviour can be categorized into three levels of strategic, tactical, and operational. The focus of this study is on the behaviour of pedestrians at operational level. Short term decisions such as acceleration and direction change behaviours are made at this level. Interactions between pedestrians (*e.g.,* Leader-follower, collision avoidance, and group behaviour) significantly affect pedestrian behaviours at operational level. According to literature, different models have been proposed to capture these interactions and study their impact on walking behaviour. Based on leader- follower behaviour, a pedestrian may follow another individual to take advantage of the space she creates. On the other hand, collision avoidance behaviour captures the tendency of pedestrians to keep distance from each other in order to avoid potential collisions. Moreover, pedestrians belonging to the same social group affect each other's walking behaviour. Approaches suggested to model pedestrian walking behaviour can generally be categorized into macroscopic and microscopic models. Microscopic models focus on the behaviour of pedestrians at individual level. Cellular automata and Social force models are the two most famous types of microscopic models.

Chapter Three: Discrete Choice Pedestrian Walking Behaviour Model

In Chapter 2, some pedestrian *modelling* approaches were reviewed. In this chapter the pedestrian walking model proposed by Robin *et al.* (2009) will be discussed into detail. Developing a model compatible with a specific application needs a deep understanding of the topic under consideration as well as the methodological and theoretical background of the model. Therefore, we begin this chapter with an introduction to general aspects of discrete choice modelling (DCM) which is used to predict pedestrians' decisions in the discussed model where walking is considered as a sequence of choices (Robin *et al.* 2009). Discrete choice theory deals with the choice behaviour of decision-makers when they have to select among a set of alternatives. In other words, it aims to model the decision process of individuals when facing a choice situation.

Discrete choice modelling has been of interest in various disciplines including econometrics and transportation engineering in recent years. McFadden (1978), McFadden (1981), McFadden &Train (2000), and Börsch-Supan (1990), among others, applied DCM in different econometrics fields such as marketing and finance. It also has been used in transportation science with application to model, among others, route choice, travel mode choice, departure time choice, and destination choice behaviour (*e.g.,* Ben-Akiva & Lerman 1985, Ben-Akiva & Bierlaire 1999,

Ben-Akiva & Bierlaire 2003, Vovsha 1997, Vovsha & Bekhor 1998 , Koppelman & Sethi 2005, Hess 2005, Lemp *et al.* 2010, Hess *et al.* 2012, and Wang *et al.* 2008).

This chapter is structured as follows: The theoretical and practical features of discrete choice models are briefly overviewed in section 3.1 and then, in section 3.2 the pedestrian walking model suggested by Robin *et al.* (2009) will be reviewed. This model focuses on the short range walking behaviour of pedestrians, while the higher levels of behaviour, namely the destination and the route are considered to be known.

3.1 Discrete Choice Models

An individual's behaviour when choosing among a set of alternatives can be predicted by means of discrete choice modelling. They are disaggregate models since they represent the choice behaviour at the level of *individual* decision-makers.

In the framework of discrete choice model some assumptions regarding choice behaviour should be taken into account. These assumptions can briefly be explained as follows: Decision-maker is supposed to be an 'individual'; it does not necessarily mean that a decision-maker has to be only one person. But instead, decision-maker can be a group of people, such as an organization or a family whose decisions would be considered as a whole and the internal interactions between them would be disregarded (Ben-Akiva & Lerman 1985). Individuals have heterogeneous tastes and preferences in choice situations. Accordingly, the socio-economic characteristics of the decision-maker such as age, gender, education and income should be considered in the model to examine the effects of these variables on choice behaviour (Ben-Akiva & Bierlaire 1999).

The set of all alternatives that are known and available to the individual during the process of decision making is called the choice set. Since here the choice *modelling* is studied in the discrete phase, the number of alternatives is finite and is referred to as the discrete choice set. The process of defining the list of alternatives is called choice set generation. In situations that the decision-maker is not aware of the alternative, the probabilistic choice set generation models can be presented to address this uncertainty (Ben-Akiva & Bierlaire 1999). Each alternative has a set of attributes which describe its attractiveness. Alternative-specific attributes are those attributes that are assigned to one specific alternative and are not common in all alternatives. Sometimes an attribute is a function of measured data instead of being directly measurable from the available data set.

The decision rule is the process performed by the decision-maker to assess available alternatives and come to a decision to select a unique choice. In general, discrete choice models are derived based on the utility theory and under an assumption of utility maximization (Ben-Akiva & Lerman 1985). In this theory, Decision-makers assign a value, called utility, to each alternative. Utility is the index of attractiveness of an alternative. Decision-makers use the concept of trade-offs to compare different attributes and are assumed to select an alternative that provides the highest utility among all the available alternatives. Some inconsistencies have been observed in the discussed approach in decision-making experiments. Therefore, the complication of human choice behaviour suggests including a probabilistic dimension in the model (Ben-Akiva & Bierlaire 1999).

The model is composed of utility functions that are consisted of observable variables and unknown parameters. The values of unknown parameters are estimated using a dataset including observed choices of individuals gathered in a similar choice situation.

In sub-section 3.1.1 Random utility models are discussed. Then in the following sub-sections, some different types of discrete choice models are explained. And finally in the last part of this section, the method used to estimate discrete choice models are briefly discussed.

3.1.1 Random Utility Models

Random utility models (RUM) are based on deterministic decision rules. The decision process is described by the concept of 'rational behaviour'. it means the decision-maker has consistent and transitive preferences while selecting among alternatives (Ben-Akiva & Lerman 1985). Consistent preference means the decision-maker repeats the same choice in similar situations. Transitive preference suggest that if the decision-maker prefers alternative one over alternative two and also prefers alternative two to alternative three, therefore alternative one is preferred to alternative three. It is assumed that while the decision-maker can perfectly discriminate between alternatives, the analyst's lack of information and limited ability to observe the decision process bring about uncertainty and therefore utilities are treated as random variables to capture this uncertainty (Ben-Akiva & Bierlaire 1999). Four sources of uncertainty have been identified by Manski (1977). It was suggested that the uncertainty exists as a result of unobserved attributes of alternative, unobserved individual's socio-economic characteristics (unobserved taste variations), measurement errors, or the use of instrumental variables.

Given a choice set *Cn*, including *J* alternatives, the utility function, *Uin,* that decision-maker *n* perceives for alternative *i*, where *i* = 1*, ..., J,* is given by:

$$U_{in} = V_{in} + \varepsilon_{in}, \qquad \text{(Equation 3. 1)}$$

where, V_{in} is the deterministic (non-random) part and is called systematic (or representative) component of the utility. It represents that part of the utility that can be observed by the analyst. Since there are some aspects of utility that are not known to the analyst, utility is also constituted of a random component to capture the uncertainty which is represented by ε_{in}. The distribution of ε_{in} is defined based on the analyst's specification of the choice situation. Various choice models are derived from different assumptions about the joint distribution of the random vector $\varepsilon' = (\varepsilon_{1n}, \varepsilon_{2n}, ..., \varepsilon_{jn})$ (Train 2003; Ben-Akiva & Lerman 1985). The alternative that provides the highest utility is chosen. Therefore, the probability that the decision-maker *n* selects alternative *i* in the choice set *Cn* is:

$$\begin{aligned} P_n(i|C_n) &= P_n(U_{in} \geq U_{jn}, \forall j \in C_n, j \neq i) \\ &= P_n(V_{in} + \varepsilon_{in} \geq V_{jn} + \varepsilon_{jn}, \forall j \in C_n, j \neq i) \\ &= P_n(\varepsilon_{jn} - \varepsilon_{in} \leq V_{in} - V_{jn}, \forall j \in C_n, j \neq i) \end{aligned} \qquad \text{(Equation 3.2)}$$

From the above equation, it can be concluded that only the difference between the utilities affects the choice probabilities and is relevant in decision process, not the absolute value of the utility. In other words, the concept of utility is relative. Considering this fact, it can be noted that only parameters that represent the differences between alternatives can be estimated in the model (Train 2003). The utility has no unit and the scale of the utility does not affect the decision process. Adding the utility of all alternatives by a constant or multiplying the utilities by a positive constant does not change the individual's choice. Therefore, the scale of the utility has to be normalized and consequently scale parameter (μ) has to be selected. The scale of the utility and variance of the unobserved portion of the utility are related. If utility is multiplied by λ, the

variance of the error term will be multiplied by λ^2 (Ben-Akiva & Lerman 1985) . Therefore, normalization is usually done on the basis of the normalization of the variance of the error terms (random terms).

Usually, the deterministic part of the utility function also contains a constant which is specific to the alternative. If we consider $m_i = E[\varepsilon_{in}]$ as the mean of the error term, we have a new random variable $e_{in} = \varepsilon_{in} - m_i + c$ in which $E[e_{in}] = c$. therefore, the deterministic part of the utility will be $V_{in} + m_i$ and the random part will be e_{in} where m_i is an Alternative Specific Constant (ASC) (Ben-Akiva & Bierlaire 1999). The alternative-specific constant represents the mean of the random term of the alternative and captures the average effect of unobserved factors on the utility. Only the differences between the alternative-specific constants are important since only differences between utilities are relevant.

The systematic part of the utility is a function of the attributes of the alternative and the socio-economic characteristics of the decision-maker. The attributes of the alternative i perceived by the decision-maker n is denoted as Z_{in} and S_n is the characteristics of the decision-maker n. Function h is defined to create a new vector from both Z_{in} and S_n to be an appropriate representation of a specific application. That is: $x_{in} = h\ (Z_{in},\ S_n)$. Therefore, the systematic component of the utility is defined as: $V_{in} = V\ (x_{in})$. If the utility is specified as linear in the parameter function, it has the following form:

$$V_{in} = \sum_k \beta_k x_{ink} \qquad \text{(Equation 3. 3)}$$

Where the vector of parameters β has to be estimated and contains the coefficients of the attributes that are included in the model. For the sake of interpretation, it is important to note that

each estimated coefficient is scaled by the scale parameter and reflects the impact of the observed factor relative to the variance of the unobserved factors (Train 2003). The variance of unobserved factors for different decision-makers may not be the same. For example, for different datasets usually different variance for unobserved factors is expected, resulting in different scale parameters for each dataset. It should be noted that the behaviour of interest may not always be totally explainable by linear parameters. Therefore, nonlinear formulation is sometimes used to represent some parameters in utility function.

Random part of the utility can be presented by different models. Logit and Probit models are the most popular ones for this purpose. In Probit models, unobserved parts of the utility are normally distributed. These models are able to capture all correlations among alternatives. But, due to their complex formulation they are not used as often as Logit models (Ben-Akiva & Bierlaire 1999). Some of the most used models from Logit family are discribed in the following.

3.1.2 Multinomial Logit Model

Multinomial logit (MNL) model is the most widely used form of discrete choice models. The closed form mathematical structure of choice probabilities makes the estimation and interpretation easy in this model. Error terms are assumed to be independently and identically type I extreme value (Gumbel) distributed (Ben-Akiva & Bierlaire 1999). The density and the cumulative distribution of each error term are defined as follows:

$$\begin{aligned} F(\varepsilon_{in}) &= e^{-e^{-\mu(\varepsilon_{in}-\eta)}}, \quad \mu \succ \\ f(\varepsilon_{in}) &= \mu e^{-\mu(\varepsilon_{in}-\eta)} e^{-e^{-\mu(\varepsilon_{in}-\eta)}} \end{aligned} \qquad \text{(Equation 3. 4)}$$

where η is a location parameter and μ is a positive scale parameter. The variance of this distribution is $\pi^2 / 6\mu^2$ which is used to normalize the scale of the utility.

The probability that decision-maker *n* chooses alternative *i* within the choice set *Cn* is defined by:

$$p(i \mid c_n) = e^{\mu V_{in}} / \sum_{j \in C_n} e^{\mu V_{jn}} \qquad \text{(Equation 3. 5)}$$

where the location parameter is set to zero. MNL model has some limitations which restrict its power of predictability and encourage the analyst to employ more flexible models. These limitations are described briefly in the following:

One of the most important properties of MNL model is the Independence from Irrelevant Alternatives (IIA). This property can be described with the following statement: the ratio of the choice probabilities of any two alternatives *i* and *j* does not depend on systematic utilities of any other alternatives. That is, selecting between *i* and *j* is independent of what other alternatives are available besides these two alternatives and what their attributes are. In other words, the ratio of the logit probabilities of any two alternatives is not affected by the choice set (Ben-Akiva & Bierlaire 1999; Train 2003). Suppose choice sets C_1 and C_2 are both subsets of choice set C_n ($C_1 \subseteq C_n$ and $C_2 \subseteq C_n$), then for any two alternatives *i* and *j* in both C_1 and C_2, IIA property can be shown as follows:

$$p(i \mid c_1) / p(j \mid c_1) = p(i \mid c_2) / p(j \mid c_2) \qquad \text{(Equation 3. 6)}$$

IIA property is not realistic in all choice situations and will bring about limitations in some cases. Some attributes cannot be observed by the analyst. Therefore, two alternatives may share some similar unobserved attributes, which results in correlation patterns between alternatives in the choice set. This correlation cannot be captured by MNL model due to IIA property.

Another limitation of MNL relates to taste variations. Decision-makers' choice behaviour varies across the population. Decision-makers have different tastes and they generally do not assign the same value to the attributes of the alternatives. In addition to the observed socio-economic characteristics, there are unobserved variables and different individual preferences which bring about these differences in decision-makers' behaviour in choice situations. Just owing to the fact that different persons are different, individuals may act differently and make different choices (Train 2003). MNL model can only capture the systematic taste variation related to observed variables. The deterministic taste variations can be represented by linear-in-parameters or non-linear formulation in the model. Tastes that vary based on unobserved factors or on the random basis cannot be captured by this model. To include random taste variation, more flexible models have to be used. That is, parameters that have to be estimated are specified to be randomly distributed across the population which is mathematically corresponding to the error component specification other than that is defined for MNL model (independently and identically distributed). The interested reader is referred to (McFadden & Train 2000; Train 2003) for more details on this matter. Moreover, the sequence of choices made by the same decision-maker over time can be represented using panel data. However, unobserved factors are assumed to be independent over time for each decision-maker in MNL models. Therefore, situations in which unobserved factors that influence individuals are correlated over time and are dependent in repeated choice situations cannot be captured by MNL models (Train 2003).

3.1.3 Generalized Extreme Value Models

Generalized extreme value (GEV) models have been developed by McFadden (1978). They comprise a large family of models which can present various choice situations. Different correlation structures in the choice set resulting from shared unobserved attributes between alternatives can be captured with this family of models, giving rise to different kinds of models such as MNL, nested logit (NL), and cross-nested logit (CNL) models (Train 2003; Ben-Akiva & Lerman 1985). The unobserved portions of utility in GEV models are described by a multivariate extreme value distribution. Choice probabilities generally take a closed form. The cumulative distribution function of error terms are defined as follows:

$$F_{\varepsilon}\ (V_1,\ldots,V_J) = e^{-G(e^{-V_1},\ldots,e^{-V_J})} \qquad \text{(Equation 3. 7)}$$

where J is the number of alternatives in the choice set C_n . Let's denote $exp\ (V_j)$ for each j as y_j, then, the function G which depends on y_j for all j (that is $G= G\ (y_1, \ldots, y_J)\ )$ is a differentiable function which has certain characteristics as follows:

1. G is non-negative ($G \geq 0$) for all $y \in R^J_+$

2. G is homogeneous of degree $\mu > 0$, That is, if each y_j is multiplied by some amount ρ, G rises by amunt ρ^{μ}: $G\ (\rho y_1, \ldots, \rho y_J) = \rho^{\mu}\ G(y_1,, \ldots, y_J)$

3. $\lim_{y_i \to \infty} G(y_1, y_2 \ldots, y_J) = +\infty, \text{for each } i = 1, 2, \ldots, J$

4. the kth partial derivative of G with regard to k different y_i, is non-positive if k is even and non-negative if k is odd, i.e. for any distinct indices $i_1, \ldots, i_k \in \{1, \ldots, J\}$:

$$(-1)^k \frac{\partial^k G}{\partial y_{i_1} \ldots \partial y_{i_k}}(y) \leq 0, \forall y \in R^J_+ \qquad \text{(Equation 3. 8)}$$

If G meets these requirements and G_i be the derivative of G with respect to y_i, that is, $\partial G / \partial y_i = G_i(y_1,...,y_J)$, $i=1,..., J$, the choice probability of choosing alternative i by the given decision-maker is:

$$P(i \mid c) = \frac{y_i G_i(y_1,...,y_J)}{\mu G(y_1,...,y_J)} \qquad \text{(Equation 3. 9)}$$

This general formulation defines the GEV family of models. Different models can be derived from proper specification of function G. For example, with defining G as $\sum_{i=1}^{J} y_i^{\mu}$ the GEV model represents MNL model (Ben-Akiva & Bierlaire 1999).

3.1.4 Nested Logit Model

Nested logit (NL) Model, first suggested by Ben-Akiva (1974) is an extension of MNL Model where the alternatives between which the decision-maker has to choose can be divided into subsets known as nests (Ben-Akiva & Bierlaire 1999). This type of model allows for some particular patterns of correlation between alternatives. It is assumed that IIA property holds for any two alternatives within the same nest, that is, the ratio of probabilities of the two alternatives is independent of the existence of any other alternative in that nest. On the other hand, IIA does not hold for alternatives in different nests, i.e. the ratio of probabilities can be influenced by the attributes of other alternatives in two nests (Train 2003; Ben-Akiva & Bierlaire 1999). The choice set Cn is partitioned into M nests, (C_{mn}), in the way that:

$$C_n = \bigcup \qquad \text{(Equation 3. 10)}$$

and,

$$C_{mn} \cap C_{m'n} = \phi \quad \forall m \neq m' \qquad \text{(Equation 3. 11)}$$

The vector of unobserved portion of utility, $\varepsilon = (\varepsilon_{1n}, \ \varepsilon_{2n}, ..., \varepsilon_{jn})$, has the GEV cumulative distribution function. The utility of alternative i belonging to nest C_{mn} can be presented by the following equation:

$$U_{in} = \tilde{V} \qquad \tilde{} \qquad \tilde{} \qquad \text{(Equation 3. 12)}$$

This equation includes terms specific to the nest which are dependent on variables that only vary over nests but not over alternatives within each nest. It also includes those terms related to the alternative itself that depend on variables which describe the alternative and differ over alternatives within nest C_{mn} (Train 2003). ε_{in} and ε_{cmn} are independent. Error terms ε_{in} are independent and identically Gumbel distributed with scale parameter μ_m which can be different for each nest, indicating different correlation between unobserved factors within each nest ($0 < \mu_m \leq 1$). The ε_{cmn} is distributed in the way that the random term $\max_{j \in c_{mn}} U_{jn}$ is Gumbel distributed with scale parameter μ. Each nest inside the choice set has a *composite utility*, defined as follows:

$$V_{C_{mn}} = \tilde{V} \quad \mu_m \quad \ln \sum_{j \in C_{mn}} e^{\mu_m \tilde{V}} \qquad \text{(Equation 3. 13)}$$

The second component in the above equation is termed *expected maximum utility*, *inclusive value* or *accessibility* in the literature and presents the utility of the best alternative in a subset (nest) of choices as the value the decision-maker assigns to that nest. In other words, it is the expected

utility that the decision-maker obtains from the choice among the alternatives in the nest C_{mn} (Ben-Akiva & Lerman 1985; Train 2003).

With the above specification, the probability of nested logit can be defined as the product of two standard logit probabilities, namely, the probability that an alternative within nest C_{mn} is selected and the probability that the alternative i is selected given that an alternative in nest C_{mn} is selected. Therefore, the probability of choosing alternative i within the nest C_{mn} by individual n can be written as follows:

$$P(i \mid C_n) = P(C_{mn} \mid C_n) P(i \mid C_{mn}) \qquad \text{(Equation 3. 14)}$$

where

$$P(C_{mn} \mid C_n) = \frac{e^{\mu V_{C_{mn}}}}{\sum_{l=1}^{M} e^{\mu V_{C_{\ln}}}} \qquad \text{(Equation 3. 15)}$$

and,

$$P(i \mid C_{mn}) = \frac{e^{\mu_m \tilde{V}}}{\sum_{j \in C_{mn}} e^{\mu_m \tilde{V}}} \qquad \text{(Equation 3. 16)}$$

where $P(i \mid C_{mn})$ is the conditional probability of choosing alternative i given that an alternative in nest C_{mn} *is* chosen and $P(C_{mn} \mid C_n)$ is the marginal probability of choosing an alternative in nest C_{mn}. The logit formula expresses the marginal probability as if it is derived from a model describing choices among nests and likewise the conditional probability is expressed by the logit formula as if it is derived from a model describing choices among the alternatives within the

nest. The *inclusive value* connects the marginal probability and the conditional probability by conveying information from the conditional probability into the marginal probability. Roughly speaking, the probability of choosing nest C_{mn} depends on the expected utility that the decision-maker receives from that nest disregarding which alternative she selects in the nest along with the expected utility that she receives by being able to choose the best alternative in the nest, which is the inclusive value ($\frac{1}{\mu_m}\ln\sum_{j\in C_{mn}} e^{\mu_m \tilde{I}}$)(Train 2003).

Parameters μ and μ_m represent the correlation between alternatives in the nest C_{mn}. Error terms of any two alternatives within nest C_{mn} are correlated. But, for any two alternatives belonging to different nests, the error terms are not correlated (Train 2003; Ben-Akiva & Bierlaire 1999). The covariance between the utility of two alternatives *i* and *j* is:

$$COV(U_{in},U_{jn}) = \begin{cases} \text{var}(\tilde{\varepsilon} & \text{and j } \varepsilon \text{ C}_{mn} \\ 0 & \text{otherwise} \end{cases} \qquad \text{(Equation 3. 17)}$$

and the correlation between the utility of two alternatives *i* and *j* is:

$$Corr(U_{in},U_{jn}) = \begin{cases} 1-\frac{\mu^2}{\mu_m^2} & \text{if i and j } \varepsilon \text{ C}_{mn} \\ 0 & \text{otherwise} \end{cases} \qquad \text{(Equation 3. 18)}$$

Since the correlation is non-negative, the following inequality has to be satisfied:

$$0 \le \frac{\mu}{\mu_m} \le 1 \qquad \text{(Equation 3. 19)}$$

A higher value of this ratio (μ / μ_m) indicates less correlation and greater independence among the unobserved portions of the utility of alternatives. When μ / μ_m is equal to one, then there is no correlation between the utility of alternatives i and j in the nest C_{mn} (i.e. $Corr(U_{in}, U_{jn})$=0). When μ / μ_m=1 for all m, indicating choices inside all nests are independent and therefore nesting becomes unnecessary. Therefore, cumulative distribution of error terms and choice probabilities simply turn into that of MNL model.

μ and μ_m cannot be identified separately and only their ratio has an important effect on the model. Therefore, usually one of them would be set to one arbitrary value such as one. If the scale parameter μ is normalized to 1, the model is called "normalized from the top". And, if one of the parameters μ_m is set to 1, the model is called "normalized from the bottom" (Ben-Akiva & Lerman 1985; Ben-Akiva & Bierlaire 1999).

Nested logit model is a GEV model with G defined as follows:

$$G = \sum_{m=1}^{M} \left(\sum_{i \in C_{mn}} y_i^{\mu_m} \right)^{\frac{\mu}{\mu_m}} \qquad \text{(Equation 3. 20)}$$

3.1.5 Cross-Nested Logit Model

Cross-nested logit (CNL) model is an extension of nested logit model. Nested logit model can only capture choice situations where alternatives within each nest are correlated and each alternative belongs to only one nest, which would be a limitation in some situations. Sometimes, it would be useful to have a model in which the unobserved portion of the utility for alternative i

is correlated with other alternatives in the nest C_{mn} and also correlated with that of other alternatives in the nest $C_{m'n}$. In other words, alternative i is a member of two nests (Train 2003). Several types of GEV models with overlapping nests have been defined letting an alternative to be a member of more than one nest. Cross-nested logit is a case in point which was first proposed by McFadden (1978) as a special case of GEV model. Flexible correlation structures among alternatives can be captured with CNL formulation. Vovsha (1997) and Vovsha & Bekhor (1998) applied this model, among others, for mode choice and route choice, respectively.

The choice set Cn is partitioned into M nests (C_{mn}) as in nested logit model. The utility of alternative i is defined as follows:

$$U_{imn} = \tilde{V} \quad \tilde{} \quad \tilde{} \qquad \text{(Equation 3. 21)}$$

ε_{in} and ε_{cmn} are independent. Error terms ε_{in} are independent and identically Gumbel distributed, with scale parameter μ_m. The ε_{cmn} is distributed in the way that the random term $\max_{j \in Cmn} U_{jmn}$ is Gumbel distributed with scale parameter μ. The parameter α_{im}, for each alternative i and each nest m, is the degree of membership of alternative i in nest m. In fact, an alternative is allocated among the nests and may belong to some nests more than other nests. α_{im} represents the portion of alternative i assigned to the nest C_{mn}. α_{im} has to be nonnegative and $\sum_{m=1}^{M} \alpha_{jm} \succ$ .When this parameter has a value of zero, it means that the alternative is not in the nest. For convenience and ease of interpretation, usually the model is normalized by having $\sum_{m=1}^{M} \alpha_{jm} = 1, \forall j$ (Ben-Akiva & Bierlaire 2003; Train 2003).

the probability of choosing alternative *i* by individual *n* is given by:

$$P(i \mid C_n) = \sum_{m=1}^{M} P(C_{mn} \mid C_n) P(i \mid C_{mn}) \qquad \text{(Equation 3. 22)}$$

where

$$P(C_{mn} \mid C_n) = \frac{e^{\mu V_{C_{mn}}}}{\sum_{l=1}^{M} e^{\mu V_{C_{\ln}}}} \qquad \text{(Equation 3. 23)}$$

and,

$$P(i \mid C_{mn}) = \frac{\alpha_{im} e^{\mu_m \tilde{V}}}{\sum_{j \in C_{mn}} \alpha_{jm} e^{\mu_m \tilde{V}}} \qquad \text{(Equation 3. 24)}$$

and,

$$V_{C_{mn}} = \tilde{V} \quad \mu_m \quad \text{n} \sum_{j \in C_{mn}} \alpha_{jm} e^{\mu_m \tilde{V}} \qquad \text{(Equation 3. 25)}$$

CNL model is a GEV model with the following generating function which first was suggested by Ben-Akiva & Bierlaire (1999):

$$G = \sum_{m=1}^{M} \left(\sum_{j \in C_n} \alpha_{jm} y_j^{\mu_m} \right)^{\frac{\mu}{\mu_m}} \qquad \text{(Equation 3. 26)}$$

With G function specified above, the probability of choosing alternative *i* is given by:

$$P(i \mid C) = \sum_{m=1}^{M} \frac{(\sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} y_j^{\mu_m})^{\frac{\mu}{\mu_m}}}{\sum_{n=1}^{M} (\sum_{j \in C} \alpha_{jn}^{\mu_n/\mu} y_j^{\mu_n})^{\frac{\mu}{\mu_n}}} \frac{\alpha_{im}^{\mu_m/\mu} y_i^{\mu_m}}{\sum_{j \in C} \alpha_{jm}^{\mu_m/\mu} y_j^{\mu_m}} \qquad \text{(Equation 3. 27)}$$

where,

$$y_j = \exp(V_j), \ \forall j \qquad \text{(Equation 3. 28)}$$

NL model is a special case of CNL model in which α_{im} is equal to one if alternative *i* belongs to nest C_{mn} and zero otherwise.

The research conducted in this thesis applies CNL model to model pedestrian walking behaviour. For more details about the derivation of other models of GEV family, the interested reader is referred to Ben-Akiva & Lerman (1985), Train (2003), and Wen & Koppelman (2001).

3.1.6 Model Estimation

The estimation process of discrete choice models involves maximizing the likelihood function. Log-Likelihood function is specified by the following general form:

$$LL\ (\beta) = \sum_{n=1}^{N} ln\ P_n(\beta)/N \qquad \text{(Equation 3. 29)}$$

Where $P_n(\beta)$ is the probability of the observed results for decision-maker *n,* β is the vector of parameters, and N is the sample size. Maximum likelihood estimation (MLE) method estimates the parameters of the model (β) with a set of values that make the observed results the most likely (Train 2003). In other words, this method finds values for model parameters that maximize the likelihood function ($LL\ (\beta)$) and consequently maximizes the fitness of the given data to the model. Since the likelihood is the probability value in range of 0 to 1. Then log-likelihood is

always negative. Therefore, the log-likelihood value closer to zero indicates a better fitting. To find this maximum value, initial estimates for model parameters (β_0) are set and then in each following iteration, values for parameters will change to new values that make the value of likelihood function greater. This iterative process stops when the convergence criterion is met, *i.e.,* the likelihood function is maximized. Theoretically, the convergence is reached when the gradient vector of the log-likelihood, *i.e.,* the vector of first derivatives of the log-likelihood function is zero. However, in reality this quantity can have a very small value near zero but is never equal to zero. Therefore, the convergence can be justified based on 1) examining whether the components of the gradient vector are less than a value identified by the researcher; or 2) performing the statistical test of the hypothesis that all the components of the gradient vector are equal to zero (less than a critical value).

Many algorithms have been developed to maximize utility function. In this research the Biogeme package (Bierlaire 2003) is used to estimate the parameters of the proposed model. Biogeme is a free package that estimates different types of discrete choice models by employing maximum likelihood estimation method. Five different algorithms can be used for optimization process in Biogeme. All of them find the local maximum of the likelihood function (not the global maximum) and each of them has their own stopping criteria. Due to the aforementioned reasons, the answer obtained from different algorithms may differ but usually these differences are small**.**

3.2 Next Step Model

In this section, a microscopic behavioural walking model based on discrete choice modelling framework is described in detail. This model was developed by Robin *et al.* (2009) and is called 'Next Step model'. Strategic and tactical decisions and, in particular, the destination and route

choice are assumed to be exogenous to the model and known a priori. The focus is on the short range walking behaviour of pedestrians as a response to their immediate environment and to the presence of other pedestrians in normal (non-panic and non-evacuation) situations.

The discussed model defines walking as a sequence of choices where each pedestrian, at each time instant, decides which location to occupy in the next time horizon *t* among a set of possible alternatives within the area around her current position. This model and its attributes are discussed in the following sub-sections.

3.2.1 The Choice Set

Defining the space model corresponds to the definition of the choice set in this model. Unlike most of the pedestrian models (such as cellular automata) in which physical space is defined by static space model, Robin *et al.* specify a dynamic (adaptive) and individual-specific spatial discretization which has been justified by different space perceptions of different individuals. The basic structure of the space model is shown in Figure 3.1.

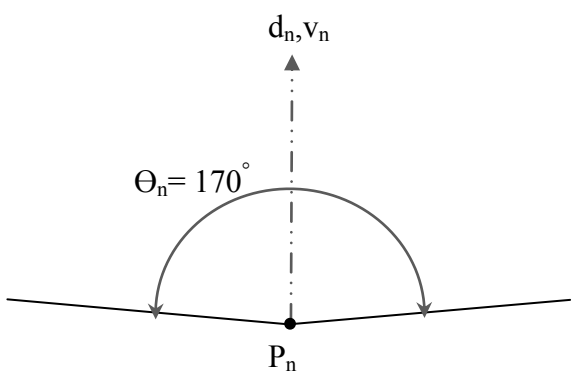


Figure 3.1: The basic elements of the space model (Source: Robin *et al.* 2009)

$p_n \equiv (x_n, y_n)$ is the current position of the decision-maker *n*, her current direction is d_n ($\| d_n \|$ =1), her current speed is v_n, and her visual angle is $\theta_n = 170°$ which generates the region of interest. It is assumed that the region of interest is located in front of the pedestrian and is overlapping with her visual field. Theoretically, the depth of decision-maker's visual field defines the size of the region of interest. And practically, if the size of the study area is reasonable, the region of interest is considered to be the same dimension of the scene floor.

A choice set consisting of 33 alternatives form all possible alternatives an individual can choose for the location she occupies in the next time step. Pedestrians walk on a two-dimensional (2-D) plane where every movement is a combination of both direction change and speed change. The Choice set is generated based on variation in decision-maker's speed and direction. Alternatives in the choice set are associated with different locations in physical space in relation to the decision-maker's current position. The current speed vector of the decision-maker determines the spatial resolution of alternatives along with the size and orientation of the choice set.

Three speed patterns and 11 predefined radial directions have been defined. With regards to the direction of movement, the visual angle is divided into 11 radial cones. The central cone represents the current direction and, consequently, the choice of not changing the direction. Ten other cones at the left and right of the central cone represent the decision of changing the direction. Uneven direction discretization is used with radial cones of smaller angular amplitudes around the current direction. The reason for that is to increase model's sensitivity to direction change with respect to the current direction. Each radial cone is characterized by its bisecting direction (*d*). Radial cones and their angular amplitudes (in degree) are depicted in Figure 3.2. Regarding speed patterns, it is assumed that the decision-maker has three options: to accelerate

up to 1.5 times her speed, to keep her current speed, or to slow down to 0.5 time her speed, creating speed zones as shown in Figure 3.3. For a time step t, the deceleration zone ranges from $0.25v_nt$ to $0.75v_nt$, the constant speed zone ranges from $0.75v_nt$ to $1.25v_nt$, and the acceleration zone ranges from $1.25v_nt$ to $1.75\ v_nt$, with their centre being at $0.5v_nt$, v_nt, and $1.5v_nt$, respectively. The authors selected the time step (t) to be 1 second.

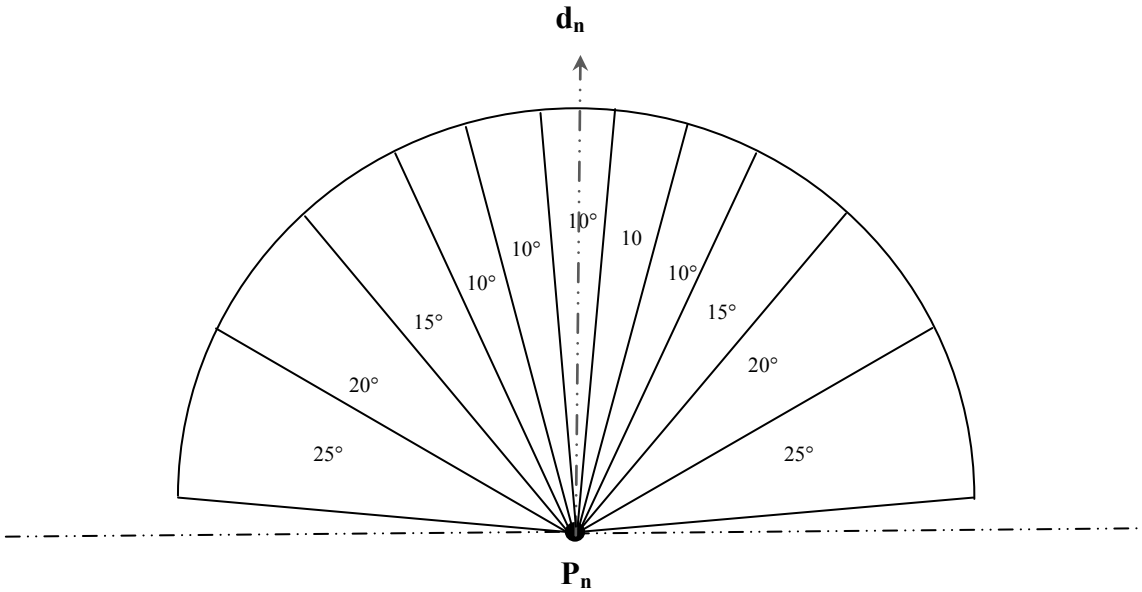


Figure 3.2: Discretization of space regarding radial directions (Source: Robin *et al.* 2009)

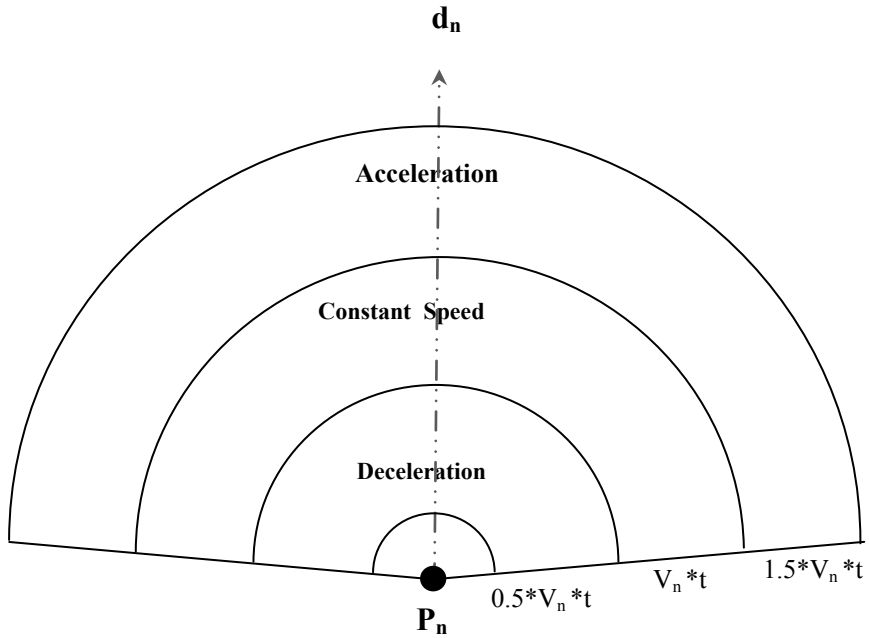


Figure 3.3: Discretization of space regarding speed patterns (Source: Robin *et al.* 2009)

Each alternative is associated with a combination of a speed pattern *v* and a radial direction *d,* generating the choice set as illustrated in Figure 3.4. Each alternative is represented by its centre which is denoted as C_{vd} . It is assumed that the individual can reach any cell's centre point in a one time step movement with sufficient change in her speed and direction. If some alternatives in the choice set are blocked with obstacles, those alternatives will be considered to be unavailable in this model. In selecting the size of the choice set (number of alternatives), the following criteria has been taken into account: First, the number of alternatives directly affects the computational efficiency of calibration process and the calculation of choice probabilities in the model. Therefore, the trade-off between having accurate spatial discretization and effective estimation process plays an important role in defining the size of the choice set. On the other hand, very fine spatial resolution would result in highly correlated alternatives where speed and direction variations do not reflect the actual choice behaviour of decision-makers. Therefore, 3 speed patterns and 11 radial directions are accepted to generate the reasonable choice set size.

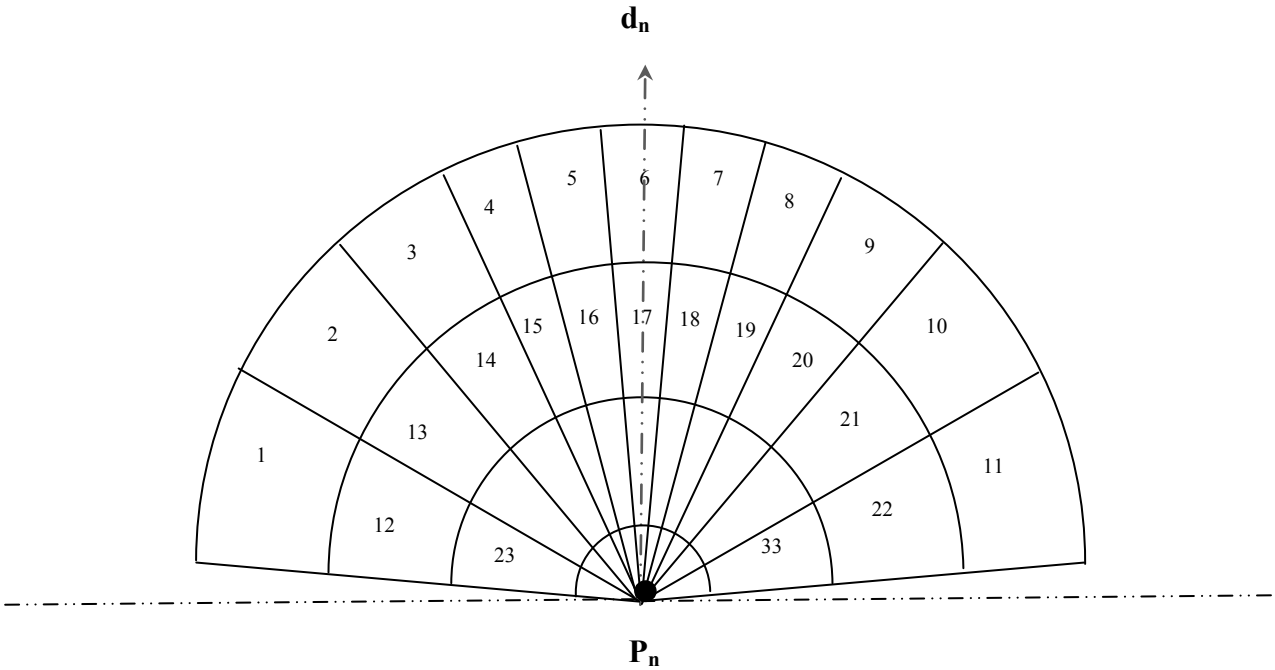


Figure 3.4: The choice set (Source: Robin *et al.* 2009)

3.2.2 The Model Specification

It is assumed that two different categories of behaviours govern pedestrian's walking. First category is the unconstrained behavioural pattern that refers to those pedestrian's behaviours that are not influenced by the presence of other individuals in the environment nearby. Second one is the constrained behavioural pattern which conversely reflects the interactions between the decision-maker and other individuals in the scene. Unconstrained behaviours are: Toward destination, Keep direction, and Free-flow acceleration. Constrained behaviours include: attractive behaviour of leader-follower and repulsive behaviour of collision avoidance. These five behaviours are assumed to control how a pedestrian performs her walking task. Therefore, the systematic part of the utility function for each alternative is composed of the terms associated with these five behaviours and with the assumption of rational behaviour, the pedestrian selects the alternative with the maximum utility.

Unconstrained Behaviours

Toward destination behaviour captures the inclination of the decision-maker to walk directly toward her destination (if it is feasible). Considering the 2-D nature of pedestrians' movements, destination influences individual's choice of speed and direction together. Therefore, the decision-maker desires to choose an alternative (location for the next time step) which minimizes the distance and the direction difference to the destination. In this model, the destination of each individual is the last location that she wants to arrive in the study area.

Keep direction reflects the tendency of the decision-maker to maintain her current direction and select an alternative that forces the minimum change in the direction of movement. This behaviour prevents from unreal and severe direction changes in pedestrians' trajectories in

normal situations. Pedestrians do not change their direction of movement regularly unless in high density situations.

Moreover, free-flow acceleration behaviour represents the tendency of the decision-maker to walk at her desired speed. However, since the desired speed is an individual specific variable which depends on socio-economic characteristics of each individual and therefore was unknown (unobserved) in this model, the authors assumed that decision-maker decides about the acceleration behaviour based on her current speed rather than the desired speed. The utility terms corresponding to each of these unconstrained behaviours are as follows:

- Toward destination

$$\beta_{ddist} ddist_{vdn} + \beta_{ddir} ddir_{dn}$$

(Equation 3. 30)

- Keep Direction

$$\beta_{dir\_central} dir_{dn} I_{d,central} + \beta_{dir\_side} dir_{dn} I_{d,side} + \beta_{dir\_extreme} dir_{dn} I_{d,extreme}$$

(Equation 3. 31)

- Free-flow acceleration

$$\beta_{dec} I_{v,dec} (v_n / v_{\max})^{\lambda_{dec}} + \beta_{accLS} I_{n,LS} I_{v,acc} (v_n / v_{\max LS})^{\lambda_{accLS}} + \beta_{accHS} I_{n,HS} I_{v,acc} (v_n / v_{\max})^{\lambda_{accHS}}$$

(Equation 3. 32)

The variable $ddir_{dn}$ is the angle between the direction of the alternative (d) and the destination in degrees. $ddist_{vdn}$ is the distance between the destination and the centre of the alternative in

metres. The two variables are illustrated in Figure 3.5. β_{ddist} and β_{ddir} were estimated with negative signs.

The variable dir_{dn} is the angle between decision-maker's current direction (d_n) and the direction of the alternative (d). Three different groups of directions have been specified for this behaviour. Alternatives belonging to each of these groups have their specific utility term. Therefore, the indicator $I_{d,central}$ is 1 if the alternative belongs to cones 5 to 7 (alternatives 5, 6, 7, 16, 17, 18, 27, 28,and 29), $I_{d,side}$ is 1 if the alternative belongs to cones3,4, 8,or 9,and finally $I_{d,extreme}$ is 1 if the alternative is in cones 1,2,10,or 11. The variable dir_{dn} is depicted in Figure 3.5. Parameters $\beta_{dir\_central}$, β_{dir_side}, and $\beta_{dir\_extreme}$ were estimated. Negative signs of all three β parameters reflect the effect of this behaviour on attractiveness of the alternative.

Regarding free-flow behaviour, two separate utility components are considered for deceleration and acceleration behaviours (the acceleration itself is composed of two terms). The first term in Equation 3.32 refers to deceleration behaviour and represents the utility of those alternatives in deceleration zone, as shown in Figure 3.3. The two other terms in this equation are related to acceleration behaviour. Indicator $I_{v,dec}$ is 1 if the alternative belongs to deceleration zone and zero otherwise. v_n is the current speed of the decision-maker. The value of reference speed is arbitrary and was set to maximum speed (v_{max} = 4.84 *m/s*). This value is fixed for all pedestrians in this model.

The utility for acceleration is divided into two parts (Second and third terms of (Equation 3. 32), one for lower speeds and one for higher speeds (v_n), in order to avoid severe reduction of utility

for high speeds. The indicator $I_{v,acc}$ is 1 if the alternative is inside the acceleration zone, as shown in Figure 3.3. $v_{\max LS}$ is the reference speed for lower speeds and has been set to 1.39 *m/s*. the indicator $I_{n,LS}$ is 1 if decision-maker's speed is lower than or equal to 1.39 *m/s* and zero otherwise. Conversely, $I_{n,HS}$ is 1 if decision-maker's speed is higher than 1.39 *m/s*. β and λ parameters were estimated; β_{dec}, β_{accLS}, β_{accHS}, and λ_{dec} have negative signs, while parameters λ_{accHs}, and λ_{accLs} are positive. It is assumed that free-flow behaviour does not affect the alternatives with constant speed (utility term for these alternatives is equal to zero). The attractiveness of accelerated and decelerated alternatives regarding this behaviour is less than the attractiveness of alternatives with constant speed which is clearly reflected in negative signs of the estimated β parameters. As the speed (v_n) increases the attractiveness of the alternative in the deceleration zone increases while the utility of alternatives belonging to acceleration zone decreases. In other words, a pedestrian who is walking at high speed has less tendency to accelerate more compared to an individual who is walking slowly.

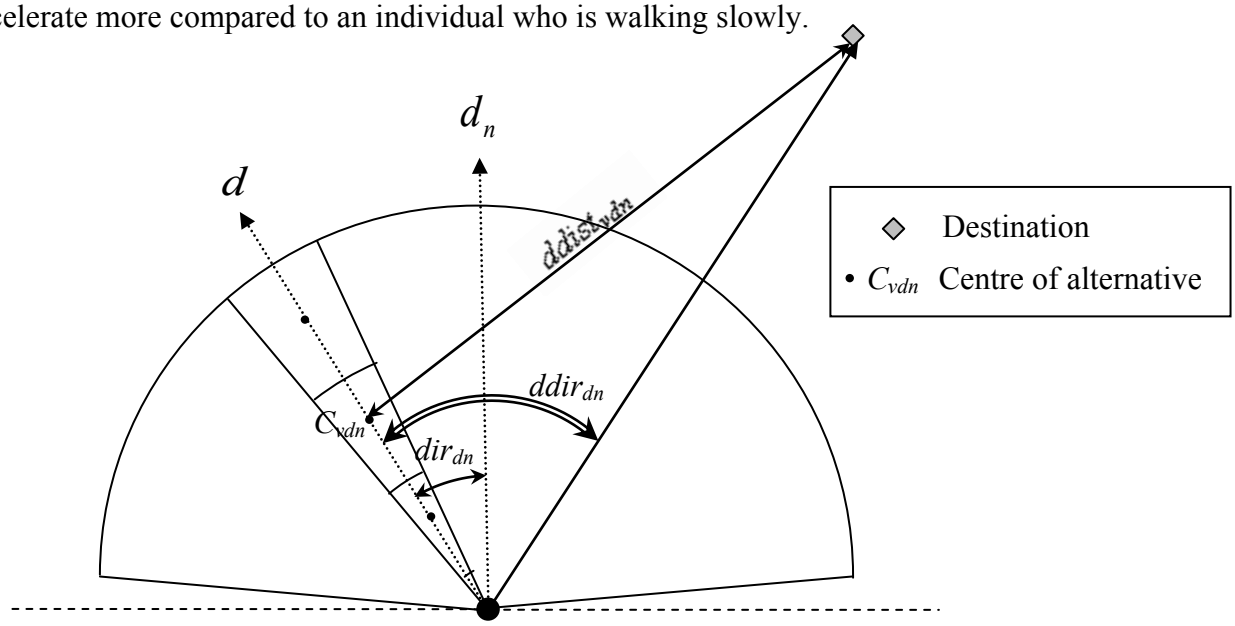


Figure 3.5: Toward destination and keep direction elements (Source: Robin *et al.* 2009)

Constrained Behaviours

Leader-follower and collision avoidance interactions are the two constrained behaviours in this model. As noted in section 2.3.1, Leader- follower is an attractive behaviour describing the pedestrian's willingness to follow another individual in order to walk more efficiently. The leader-follower behaviour was modeled using the sensitivity-stimulus framework originated from the car following behaviour in the literature. Pedestrians react to the stimuli arise from the environment. However, it was noted that vehicle movements are restricted to pre-defined lanes and facilities of the road environment and therefore, lane changing and acceleration can be modeled as two separate forms of behaviours in this case. Conversely, Pedestrians walk on a two-dimensional (2-D) plane where every movement is a combination of direction change and speed change which cannot be modeled independently. The stimulus is a function of leader's relative speed and direction. Leader is an individual in a specified area near the decision-maker whose walking direction is close to the angle representing the alternative's direction. It is assumed that the decision-maker tries to adjust her speed in response to the leader. First, a group of potential leaders are identified for each of the 11 radial cones. To be considered as a potential leader (k), a pedestrian should be in a specific region of interest. The region of interest for each cone is restricted to the left and right direction boundaries of the cone (d_l and d_r) in a way that the furthest distance in this region to decision-maker is equal to 5 times the radius of the choice set (D_{max}), as shown in Figure 3.6. The difference between a potential leader's movement direction (θ_k) and the direction of the cone where she is located (θ_d) has to be less than 10 degrees. If an individual meets all the aforementioned criteria, she will be identified as a potential leader. There may be more than 1 potential leader for each cone. Therefore, the leader is identified among the potential leaders based on her distance to the decision-maker, *i.e.,* leader

is the one with the closest distance to the decision-maker. There are two types of decelerated and accelerated leaders. If the leader's speed (v_L) is less than the speed of the decision-maker (v_n), a decelerated leader is identified. Conversely, if the leader's speed is higher than the speed of the decision-maker, an accelerated leader is identified.

Collision avoidance behaviour is a repulsive behaviour of a pedestrian who is negatively influenced by other individuals. This behaviour reflects the impact of potential collisions on walking behaviour of the pedestrian. This behaviour was also modeled by employing the sensitivity-stimulus method and was basically motivated by the concept of personal space (section 2.3.2). Collider is a pedestrian in a neighbourhood region of the decision-maker with the walking direction almost opposite to the decision-maker. It is assumed that the decision-maker changes her direction in response to the potential collision. Similar to leader-follower behaviour, first a group of potential colliders are identified for each of the 11 radial cones. A potential collider (k') is a pedestrian in a specific region of interest which is restricted to the left and right direction boundaries of the radial cone (d_l and d_r) where the alternative locates. The distance between the centre of the alternative and the potential collider should be less or equal to 10 times the radius of the choice set (D_{max}), as shown in Figure 3.7 . The difference between a potential collider's movement direction ($\theta_{k'}$) and the decision-maker's direction of movement (θ_{d_n}) has to be between 90° to 180°. If an individual meets all the aforementioned criteria, she will be identified as a potential collider. If there is more than one potential collider for the cone, the collider is identified based on the direction of movement. The decision-maker selects the collider, among all potential colliders, whose movement direction creates the largest angle with her movement direction, *i.e.,* walks in a more frontal direction. The collision avoidance behaviour even occurs at low density situations, avoiding interactions with pedestrians at long

distances. This justifies the bigger dimension of the region of interest for collision avoidance interaction compared to the leader-follower behaviour. The utility terms associated with the two constrained behaviours are as follows:

- Leader-follower

$$I_{v,acc} I^{L}_{d,acc} \underbrace{\alpha^{L}_{acc} D^{\rho^{L}_{acc}}_{L}}_{\text{Sensitivity}} \overbrace{\Delta v^{\gamma^{L}_{acc}}_{L} \Delta\theta^{\delta^{L}_{acc}}_{L}}^{\text{Stimulus}} + I_{v,dec} I^{L}_{d,dec} \underbrace{\alpha^{L}_{dec} D^{\rho^{L}_{dec}}_{L}}_{\text{Sensitivity}} \overbrace{\Delta v^{\gamma^{L}_{dec}}_{L} \Delta\theta^{\delta^{L}_{dec}}_{L}}^{\text{Stimulus}} \qquad \text{(Equation 3. 33)}$$

- Collision avoidance

$$I_{d,C} \underbrace{\alpha_{C} e^{\rho_{C} D_{C}}}_{\text{Sensitivity}} \overbrace{\Delta v^{\gamma_{C}}_{C} \Delta\theta^{\delta_{C}}_{C}}^{\text{Stimulus}} \qquad \text{(Equation 3. 34)}$$

The identifier $I_{v,acc}$ is 1 if the alternative belongs to acceleration zone and zero otherwise. $I^{L}_{d,acc}$ is 1 if the accelerated leader has been identified for the radial cone in which the alternative locates and zero otherwise. D_L is the distance between the decision-maker and the leader. $\Delta v_L = |v_L - v_n|$, which is the difference between leader's speed (v_L) and the decision-maker's current speed (v_n). $\Delta\theta_L = \theta_L - \theta_d$, which is the difference between the leader's moving direction (θ_L) and the angle corresponding to direction d (θ_d), as shown in Figure 3.6. $I_{v,dec}$ is

1 if the alternative belongs to deceleration area and zero otherwise. The value of 1 for the identifier $I^{L}_{d,dec}$ represents the presence of the decelerated leader for the cone in which the alternative locates. Parameters α^{L}_{acc}, ρ^{L}_{acc}, γ^{L}_{acc}, δ^{l}_{acc}, α^{L}_{dec}, ρ^{L}_{dec}, γ^{L}_{dec}, and δ^{l}_{dec} were estimated . As expected, α^{L}_{acc}, α^{L}_{dec}, γ^{L}_{acc}, and γ^{L}_{dec} have positive values and $\rho^{L}_{acc}, \rho^{L}_{dec}$, δ^{l}_{acc}, and δ^{l}_{dec} have negative signs to reflect the effect of the leader on the attractiveness of the alternative. Accelerated leaders increase the attractiveness of alternatives with a speed higher than the current speed of the decision-maker and decelerated leaders conversely make the alternatives in deceleration zone more attractive. Presence of the leader has no effect on alternatives with constant speed.

In Equation 3.36 the identifier $I_{d,C}$ is 1 if the collider is available for the cone in which the alternative locates and zero otherwise. D_c is the distance between the centre of the alternative and the collider. $\Delta\theta_c = \theta_c - \theta_{dn}$, which is the difference between collider's movement direction (θ_c) and decision-maker's movement direction (θ_{dn}), as shown in Figure 3.7. $\Delta v_c = v_c + v_n$, is the summation of collider's speed (v_c) and decision-maker's speed (v_n). Parameters α_C, ρ_c, γ_c, and δ_c were estimated; α_C and ρ_c have negative signs and γ_c and δ_c are positive representing that the presence of a collider in a cone decreases the attractiveness of all alternatives in that cone and the extent of this negative effect differs based on the collider's speed and direction and the distance of the collider to the alternative.

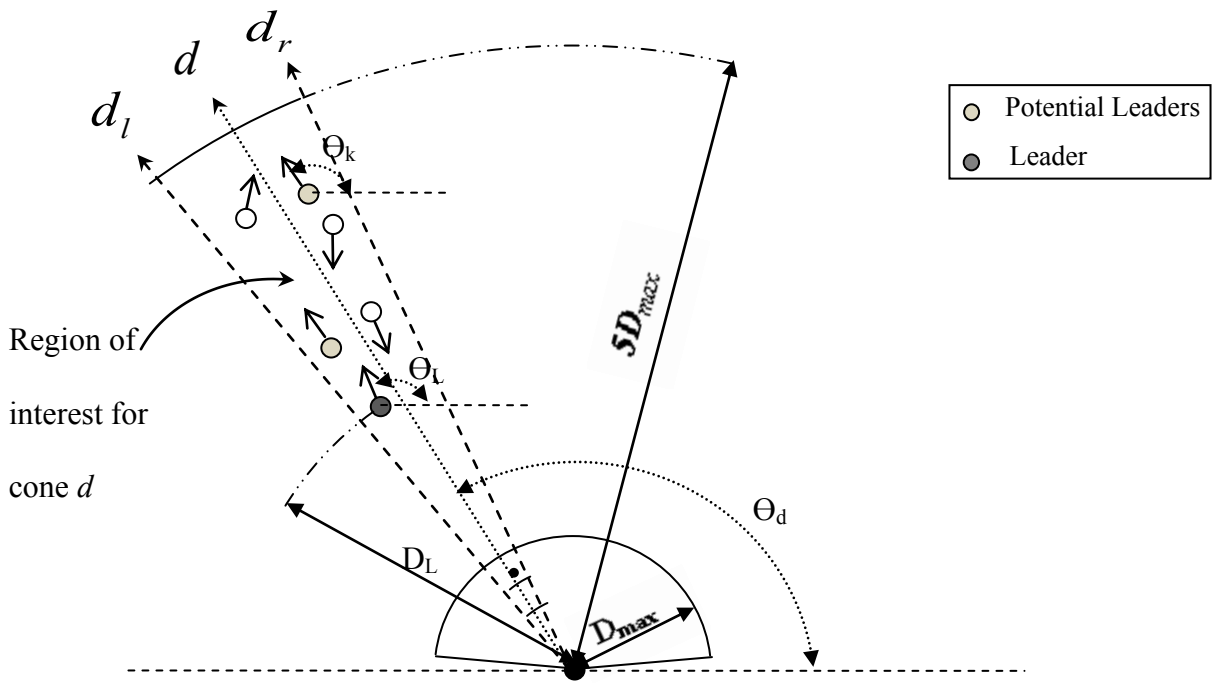


Figure 3.6: Leader identification (Source: Robin *et al.* 2009)

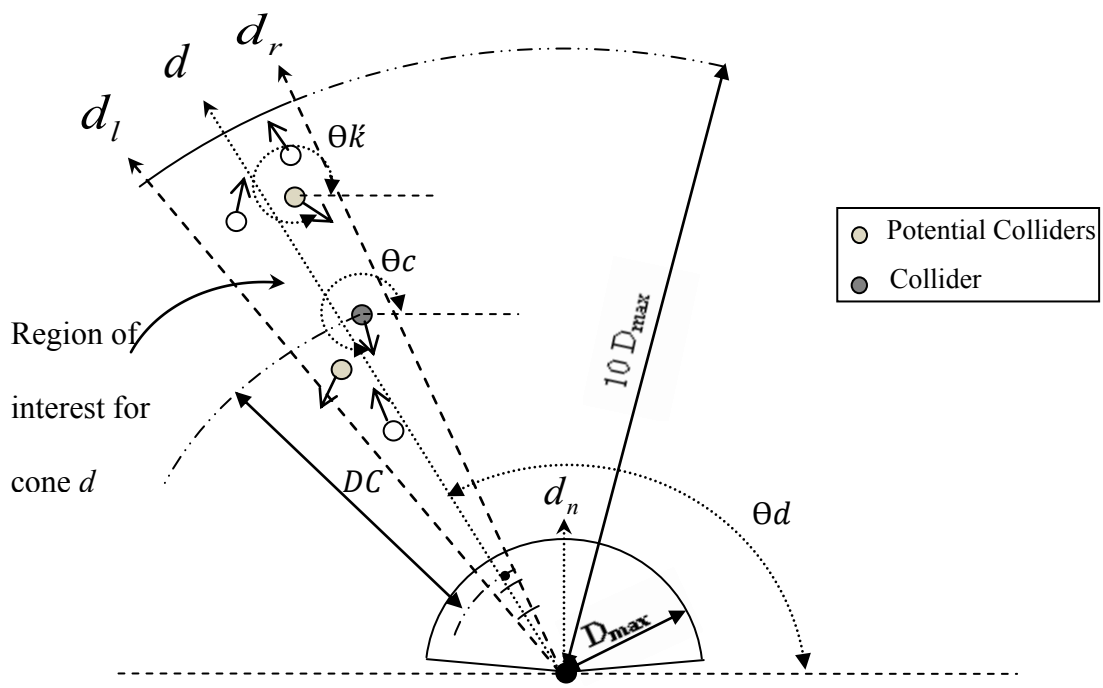


Figure 3.7: Collider identification (Source: Robin *et al.* 2009)

The model was specified based on different variables related to the attributes of alternatives. Also, only one attribute of the decision-maker (speed of pedestrian) was modeled using non-linear and deterministic specification for taste variation. Although the socio-economic characteristics of individuals affect their walking behaviour and interactions, these attributes have not been included in the model (except for speed) .Capturing the socio-economic characteristics is not an easy task and usually requires the researcher to conduct a controlled experiment. The systematic part of the utility function that decision-maker *n* perceives for alternative *i* is given by the summation of all the terms described above:

$$
\begin{aligned}
V_{in} &= \beta_{ddist} ddist_{vdn} + \beta_{ddir} ddir_{dn} \\
&+ \beta_{dir\_central} dir_{dn} I_{d,central} + \beta_{dir\_side} dir_{dn} I_{d,side} + \beta_{dir\_extreme} dir_{dn} I_{d,extreme} \\
&+ \beta_{dec} I_{v,dec} (v_n / v_{\max})^{\lambda_{dec}} + \beta_{accLS} I_{n,LS} I_{v,acc} (v_n / v_{\max LS})^{\lambda_{accLS}} + \beta_{accHS} I_{n,HS} I_{v,acc} (v_n / v_{\max})^{\lambda_{accHS}} \\
&+ I_{v,acc} I^L_{d,acc} \alpha^L_{acc} D_L^{\rho^L_{acc}} \Delta v_L^{\gamma^L_{acc}} \Delta\theta_L^{\delta^L_{acc}} + I_{v,dec} I^L_{d,dec} \alpha^L_{dec} D_L^{\rho^L_{dec}} \Delta v_L^{\gamma^L_{dec}} \Delta\theta_L^{\delta^L_{dec}} \\
&+ I_{d,C} \alpha_C e^{\rho_C D_C} \Delta v_C^{\gamma_C} \Delta\theta_C^{\delta_C}
\end{aligned}
$$

(Equation 3. 35)

Five nests have been specified based on speed and direction: Acceleration, deceleration, and constant speed (regarding speed) and central (containing alternatives 6, 17, and 28) and non-central (regarding direction). The nest structures are shown in Figure 3.8. It can be seen that each alternative belongs to two nests. Therefore, the random part of the utility function was captured with cross-nested logit (CNL) model. CNL structure captures the correlations between alternatives in different nests. The degree of membership to each nest (α_{jm}) was set to 0.5 for all alternatives. The model scale factor μ was set to 1 and scale parameters associated with the nests

(μ_m) were estimated. The probability of selecting alternative i within the choice set C is given by Equation 3.27 for CNL model.

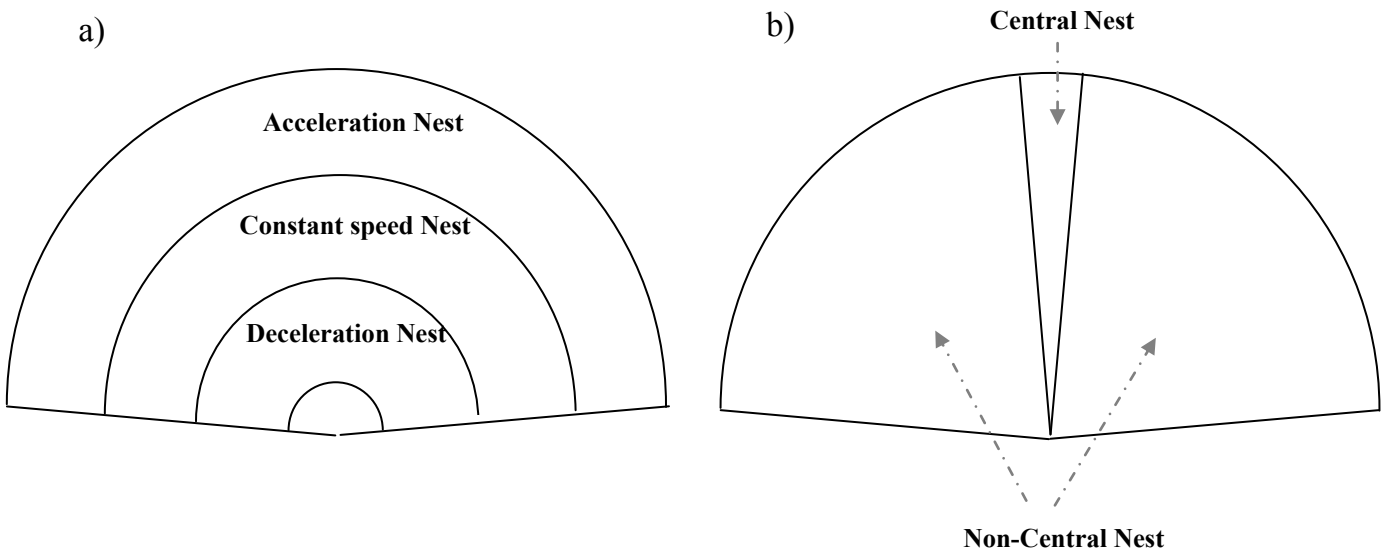


Figure 3.8: Nest structure based on a) speed b) direction

Selecting this approach to model walking behaviour in this research is justified by several reasons. Discrete choice models seem to be a proper class of models to explain pedestrian behaviours; they are disaggregate models, carried out at the level of individual decision making. Therefore, they are compatible with microscopic and agent-based approach of pedestrian *modelling* where each individual is modeled independently (Antonini 2005).

In contrast to widely used physics and force-based approaches which try to estimate pedestrian behaviours within which individuals are considered as particles, such as the popular 'social force model', discrete choice models concentrate directly on the behavioural facets of pedestrian movement where each pedestrian is a decision-maker. For example, next step model makes use of a mathematical framework relying on discrete choice analysis in which walking is considered as a sequence of pedestrians' choices. It uses econometric tools to model pedestrian behaviours

and is developed to be calibrated on real data trajectories which make it a more realistic representation of pedestrian behaviour.

The discrete choice framework comprises of sets of specific methods, including ordered, nested or cross-nested models, which contribute to its power to investigate complex behavioural processes and to deal with different patterns of correlation between alternatives. In fact, the assumption of independency between alternatives in the decision making process is not valid in all cases. For example, in next step model the multidimensional nature of the choice set results in structural correlation between alternatives which can only be captured by cross-nested or mixed models.

In addition, another valuable potential of discrete choice framework is its ability to consider and analyze the influence of individuals' different characteristics on walking behaviour. Panel data structure, expressing the presence of unobserved heterogeneity in pedestrian population, or state dependency, presenting the dependency among the subsequent choices made by the same decision-maker, can be modeled in this framework by using coefficients that vary over the population and not across the observations.

Besides the aforementioned strong points of DCM to model pedestrian walking behaviour, next step model itself is found to be robust and have a good performance in predicting choices selected by decision-makers; and it has been estimated and validated on real data. Robin *et al.* (2009) validated next step model both on the same data used to estimate the model parameters collected at a crosswalk in Sendai, Japan and the data set collected in a controlled experiment at Delft University. The estimated parameters are significant and the signs of these parameters are compatible with the assumptions.

Although next step model is a useful and valuable model in the context of pedestrian walking behaviour, just like any other model it has its own limitations.

Actual walking behaviour varies from culture to culture and in different situations (Robin *et al.* 2009). Next step model has been validated on bi-directional pedestrian flow data collected at Japan and the Netherlands. To generalize this model to different circumstances, the validity of the model in those situations should be assessed.

Model validation is an integral part of model development. Next step model has only been validated regarding decision-makers' behaviour and choices made by them, but not at macroscopic scale. That is because it has not been implemented in a simulation tool. Simulation tools enable efficient validation at both microscopic and macroscopic scales. A simulation framework would provide the possibility to evaluate model's performance regarding traffic flow characteristics such as density, flow, average speed, and egress time and also to observe pedestrians' trajectories and the way they move and validate potential self-organized patterns emerging in pedestrian crowds such as lane formations.

3.3 Summary

In this chapter, the discrete choice modelling framework and the walking behaviour model proposed by Robin *et al.* (2009) (next step model) was reviewed. Next step model explores the operational level of pedestrian walking behaviour by employing discrete choice modelling concepts. Pedestrians decide about the location they are going to occupy in the next step among the alternatives that vary depending on pedestrian's speed and direction. CNL model was used to capture the correlation structure in the alternatives' utility functions. In this model, walking behaviour of pedestrians is governed by five behavioural patterns, capturing pedestrians'

interactions with each other in addition to those behaviours that are independent from the existence of other pedestrians.

Chapter Four: Model Development for Pedestrian Group Behaviour

The purpose of the research described in this chapter is to capture the effect of pedestrian grouping on pedestrian walking behaviour using a discrete choice framework. This is achieved by extending the model presented by Robin *et al.* (2009) to account for group formation and then validating the developed model, using real data. In section 4.1 the applied methodology is explained. Data collection is described in section 4.2 and the results are reported and discussed in section 4.3. Finally, a summary is given in section 4.4.

4.1 Methodology

4.1.1 Practical Context of Group Behaviour

As described in sub-section 2.3.3, the majority of pedestrians in a crowd walk in groups instead of walking alone. The maintenance of the social bonds in a group affects individual's walking behaviour. Decisions of pedestrians in a group can be influenced by other group members. They need to balance their individual mobility while trying to maintain the group formation. Hence, this phenomenon can potentially play an important role in building a realistic walking behaviour model. However, the impact of this constrained behavioural pattern on walking behaviour has not been included in next step model proposed by Robin *et al.* (2009). The modelling of group behaviour in a discrete choice framework is a novel contribution in this thesis.

In this research, the walking behavioural patterns are extended by adding a contribution to the utility function, capturing the tendency of people in the same group to maintain group behaviour, *i.e.,* stay together and keep short distance from each other. It should be noted that the term "group" is used in its socio-psychological sense, referring to people who are socially bonded. This may arise when family members or friends walk together on purpose. Groups of pedestrians who walk close to each other by chance and without any social ties are not considered group-mates. It is assumed that pedestrians in a group continuously modify their positions to keep the group coherence and communicate easier with each other. At the same time they try to avoid collision with each other and with out-of-group nearby individuals.

4.1.2 Model Specification

The basic model elements in the proposed approach are similar to next step model in Robin *et al.* (2009). Generally, the location where a pedestrian chooses to occupy in the next time step, within the area around her current position, is modelled within a discrete choice framework. The time step is selected to be one second. The space is discretized in a dynamic and individual-specific way to build a choice set structure in front of each pedestrian at every time step. The choice set structure depends on the pedestrian's current direction and speed as shown previously in Figure 3.4. The choice set comprises of 33 alternatives, or discrete positions, in which each alternative is associated with a combination of changes in speed and direction during the next time step. The choice set is divided into three main subsets based on three speed patterns of constant speed, acceleration, and deceleration. The choice set is also divided based on the direction of movement according to 11 radial directions as illustrated earlier in Figure 3.2. With the assumption of rational behaviour, the individual selects the alternative with the maximum probability of being selected. The error structure (described in section 3.1) is captured by cross-

nested logit (CNL) model. Five nests, groups of alternatives, are specified based on speed and direction: accelerated, constant, decelerated, central and non-central. The degree of membership of an arbitrary alternative (j) to a nest (m), which is given the notation α_{jm}, is constrained to the constant value of 0.5. In addition, the model scale factor μ is fixed to 1 and other scale parameters associated with the nests (μ_{m}) are estimated.

The different components of the utility function describe two classes of movement regimes of constrained and unconstrained. Modelling group behaviour is performed by changing the model specification of components related to constrained behaviours. The components describing unconstrained walking behaviour including toward destination, keep direction, and free flow acceleration have not been changed (Equations 3.30, 3.31, and 3.32). Constrained patterns have been expanded to account for group formation and maintenance. A completely new term has been added to the utility function capturing the group maintenance behaviour. Leader-follower and collision avoidance terms have not been changed (Equations 3.33 and 3.34). Details and assumptions regarding unconstrained patterns as well as leader-follower behaviour and collision avoidance behaviour are similar to Robin *et al.* (2009) work and have been described in sub-section 3.2.2. In the following, a detailed description of newly presented group maintenance behaviour is given.

4.1.3 Utility of Group Maintenance

It is important to define some basic principles governing group behaviour based on which the utility function is specified. The simple concept of group formation is governed by the behaviour of pedestrians such that they tend to maintain the group unity while moving towards their destinations. The various intra-group structures and hierarchical relationships, *e.g.,* the leader-

follower relations within a group, are not considered in this model. The modelled behaviour is the maintenance of a coherent group while moving toward destination.

The attraction behaviour among individuals in a group is presented as an extension of discrete choice model describing pedestrian behaviour. In this research, group behaviour is described by two following group-maintenance mechanisms:

i. Each group member heads towards the centre of the group. Group centre is the average position of pedestrians in the group (excluding the decision-maker's position).

ii. Each individual in a group tends to match the direction of movement with the average moving direction of other group members.

Therefore, the reaction of an individual to other group members is captured by the following interaction terms:

$$I_g \beta_{gdist} D_g^{\rho_g} + I_g \beta_{gdir} \Delta\theta_g^{\delta_g} \qquad \text{(Equation 4. 1)}$$

where:

- I_g is equal to one if the decision-maker is a group member, and zero otherwise,
- D_g is the distance between the centre of the alternative and the projected centre of the group in the next time step, excluding the decision-maker herself. This is further illustrated in Figure 4.1.,
- $\Delta\theta_g = |\theta_g - \theta_d|$ is the difference between the direction identifying where the alternative locates and the projected average moving direction of other group members,

- β_{gdist} and β_{gdir} are sensitivity parameters for distance and direction preservation, and
- ρ_g and δ_g are parameters to capture any non-linearity in the utility contribution of deviation in distance and direction.

Regarding this behaviour, it is assumed that the decision-maker predicts the future positions of her group-mates each time step and selects an alternative which minimizes both the distance to the projected centre of the group (D_g) and the difference between her movement direction (θ_d) and the predicted average moving direction of group members, excluding decision-maker herself (θ_g). The movement direction of decision-maker in next step is the angle characterizing the radial cone where the alternative locates. The decision-maker predicts each group member's future position and direction based on the assumption of constant velocity. It means the projected position of every pedestrian in next step is expected to be cell 17 (see Figure 3.4) which is the most frequently selected choice in the model. The predicted coordinates of the group centre are taken as the average of the x-coordinates and y-coordinates of predicted positions of group members. Since cell 17 represents the choice of no change in movement direction, projected average movement direction of individuals in the group is equal to the average of current moving direction of group members. Parameters β_{gdist}, β_{gdir}, ρ_g, and δ_g have to be estimated. Negative signs are expected for β_{gdist} and β_{gdir}, and positive signs are expected for ρ_g and δ_g.

Figure 4.1 illustrates the components of this behaviour by demonstrating four pedestrians walking in a group. In this figure, the black circle represents the decision-maker and the grey circles display the decision-maker's group-mates (arrows show their directions of movement). Current movement direction of the decision-maker is denoted as d_n. Light blue circles are projected positions of group-mates in the next step and the diamond-shaped point indicates the

projected centre of the group (from decision-maker's point of view). The iterative process of model specification, calibration, and validation is performed to obtain the best result. The utility term for group behaviour has been selected after performing many trials and rejecting some other potential specifications in this iterative process. The best result is the solution that fits the given dataset better and also satisfies the considered assumptions regarding the behaviour leading to a realistic walking behaviour. For example, in an initial specification each decision-maker in a group was only linked to one other pedestrian in that group (closest pedestrian to decision-maker) and was assumed to match her distance and direction of movement with that single pedestrian. Implementing the model in a simulation tool led to rejection of this specification. This assumption caused pedestrians to constantly change their positions and consequently unrealistic trajectories were generated.

Finally, the utility term for group behaviour (presented in Equation 4.1) has been added to the utility function shown earlier in Equation 3.35.

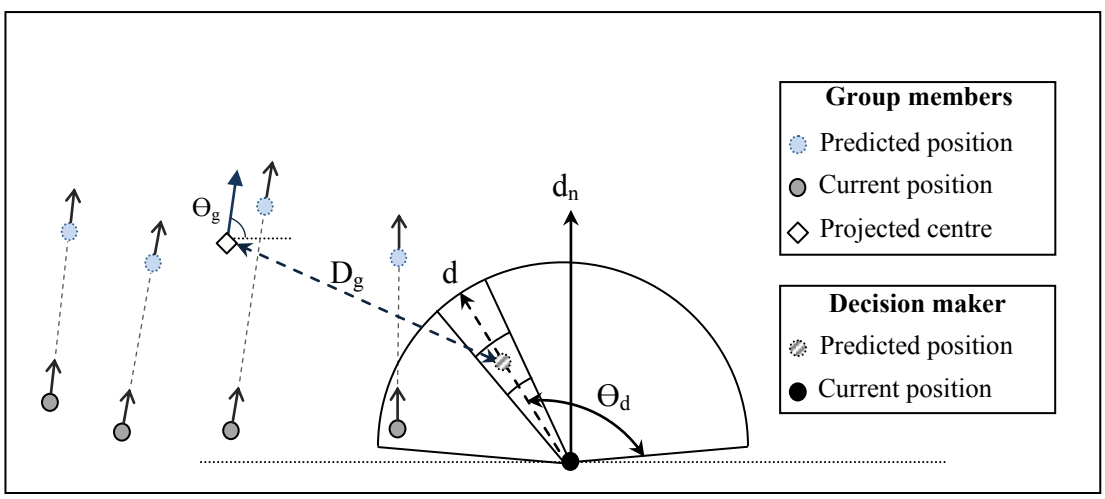


Figure 4.1: Elements of Group Behaviour

4.2 Data Collection

Data of naturalistic pedestrian movement is required in order to calibrate the model. Data was collected from actual pedestrian movements in the Downtown area of Vancouver, British Columbia at the intersection of Robson and Broughton streets (Ismail *et al.* 2009). The video was recorded during late afternoon capturing the pedestrian movements going to and coming back from a large firework event happening at the same time in the nearby area and therefore contains high pedestrian volumes. The video was recorded from the 29th floor of a building that faces the mentioned intersection. A frame extracted from this video is shown in Figure 4.2. The following sub-sections describe the process of data collection which involves the extraction of pedestrian positions and their choices.

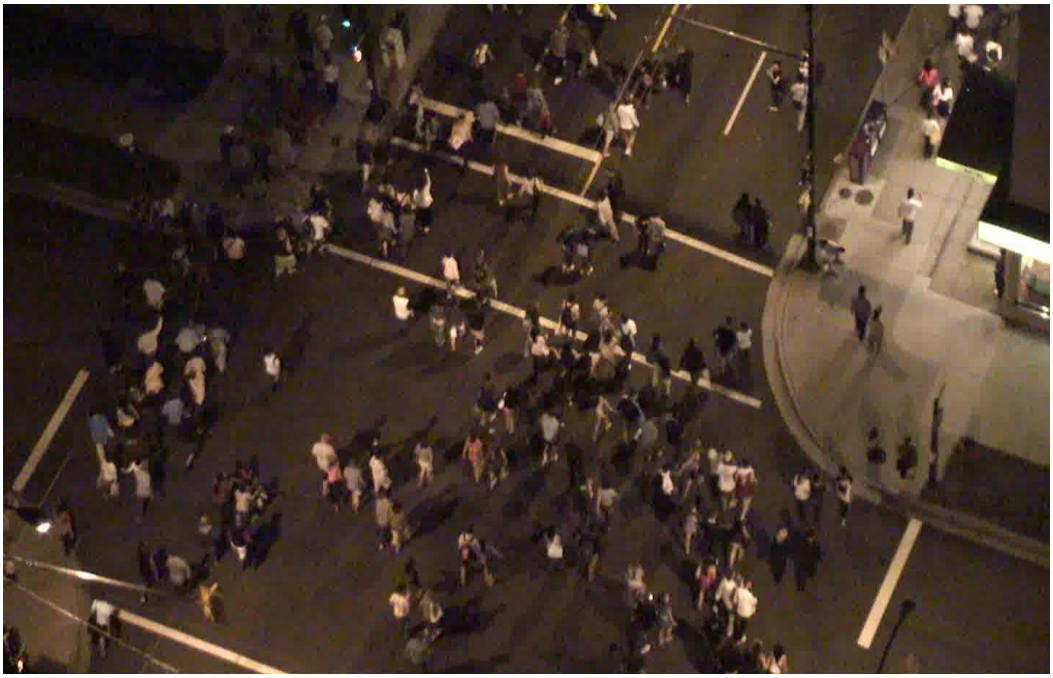


Figure 4.2: A frame from the firework data

4.2.1 Tracking of Pedestrian Positions

Larger portion of pedestrians in this crowd were walking in groups rather than walking alone. Groups of two to five people were observed in this crowd (the majority of them consisted of two people). Video sequences of 100 seconds long were selected to extract microscopic data from pedestrian movements. The video was recorded at the rate of 30 frames per second. Trajectories of pedestrians at crosswalks and along a main street, which was closed to vehicle traffic, were extracted from these video sequences, using a manual pedestrian tracking method. Centres of pedestrians' heads were tracked. Data was extracted with a time interval of 15 frames (0.5 seconds). The tracking program was written using C++ language and Qt libraries. Once a pedestrian appears in the scene the person who tracks pedestrians (tracker) draws a circle object on the centre of her head on the transparent overlay which is displayed on top of the video image and then the tracker moves this object with the movement of the pedestrian in subsequent frames (rate of 2 frames each second) until she leaves the study area. As a result, the pixel coordinates of all pedestrians' positions are stored every half a second. One of the concerns associated with manual tracking process is the difficulty to track the same point in subsequent frames which may bring about some minor errors in this process. However, the tracks were screened and the tracking process was repeated several times in order to mitigate these errors.

4.2.2 Conversion from Pixel to World Coordinates

Camera calibration was performed to estimate camera parameters using the method described in (Ismail *et al.* 2013). Camera calibration is conducted in order to project the coordinates of each observation from the image planes to real world coordinates and generate pedestrian trajectories. For this purpose, Corresponding points in image and world spaces and also some distance and

angular constraints in image plane were annotated. The distances of line segments were measured directly in the field. Angular constraints include parallel or perpendicular road markings, and parallel poles beside the road (Ismail *et al.* 2013). Figure 4.3 shows the preciseness of the estimation of camera calibration parameters by demonstrating the reference grids in both image and world space. A full description of the calibration process is out of the scope of this thesis and the interested reader is referred to Ismail *et al.* (2013). Since in this study it is assumed that the location of each pedestrian is the centre of her head, the measured distance between pedestrians with different heights is different from their actual distance due to the perspective error. However, due to the high altitude of the camera in this scene (camera position is estimated to be 88 metres above the ground), the parallax error resulted from pedestrian height difference is negligible.

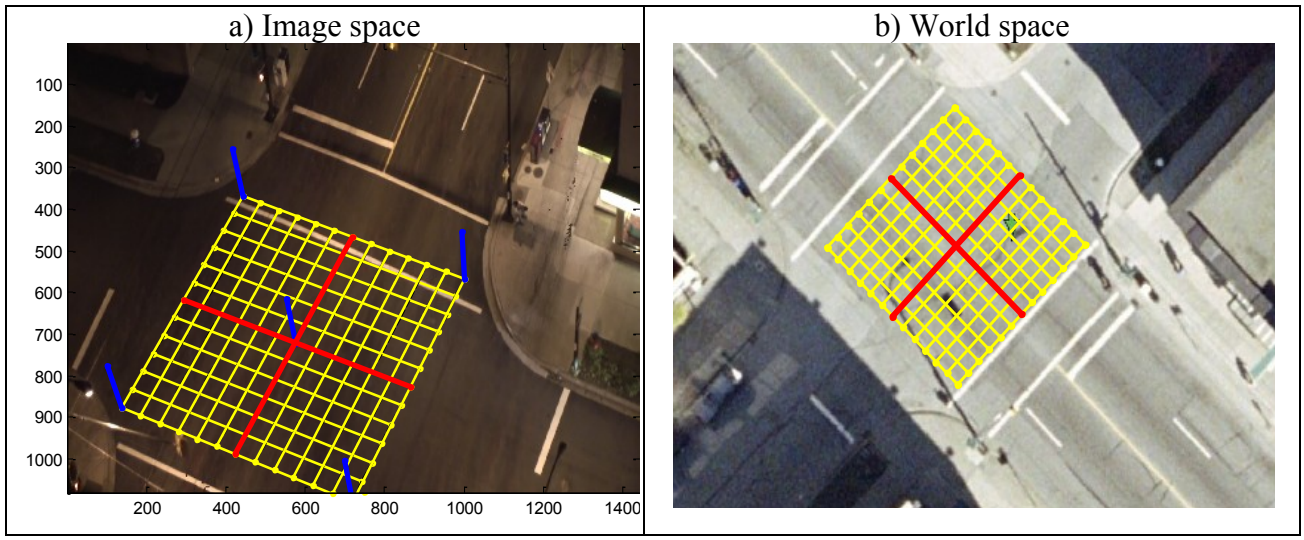


Figure 4.3: Grid references in image and world planes (Ismail *et al.* 2013)

4.2.3 Calculation of Motion Parameters

The speed and direction data for each pedestrian was derived from the successive observed positions of the pedestrian in the walking plane. Let P_t be the position of the pedestrian in time t, such that $P_t = (x_t, y_t)$, the direction and speed were calculated based on the following set of equations:

$$D_t = P_t - P_{t-0.5}$$

$$dir_t = D_t / \|D_t\|$$

$$V_t = \|D_t\| / 0.5 \qquad \text{(Equation 4. 2)}$$

where D_t is the displacement of the pedestrian during two successive observations, dir_t and V_t are direction vector and the speed of the pedestrian in time t respectively.

Observations were made every half a second (time step = 0.5sec). The term dir_t represents the pedestrian's vector of movement direction in global coordinate system. This direction is converted to local coordinates where the current direction of the pedestrian, *i.e.,* movement direction in time step $t - 0.5$, is the *y* axis of the local coordinate system. A local direction of 90° indicates no change in pedestrian's movement direction. In other words, it shows that the pedestrian's centreline of movement for that observation overlaps the local *y* axis. Negative local angles (larger than 180°) were removed from the data. These angles are either caused by some errors in manual tracking or may represent some rare situations when the pedestrian walked backwards.

4.2.4 Data Screening and Review

After extracting data from pedestrian trajectories, data has been checked for errors and the trajectories of some pedestrians have been tracked again and reviewed to remove possible errors in tracking. Criteria used to identify observations with potential errors are as follows:

Criterion 1: Observations with a speed larger than 3 m/s have been reviewed. These high values of speed are not likely to happen very often unless the pedestrian is running; however, in normal (non-panic) situations like the one presented in this data, few of them is likely to be found unless the tracker has unintentionally forgotten to scan the position of the pedestrian in a time step and as a result the reported speed is actually the representative of pedestrian's displacement in two time steps in a row by mistake. A characteristic of this error is that these high speeds are generally followed by a zero speed (pedestrian's position has not been stored in the last time step).

Criterion 2: Observations with the speed of zero have been reviewed to make sure that this occurrence is not due to the tracker's failure to move the circle object on top of pedestrian's head in that time step and the pedestrian has actually been standing in the scene.

Criterion 3: If there are many observations in a frame with the speed of zero, it may be due to the fact that the tracker has mistakenly skipped that frame. Therefore, all observations in that frame have been reviewed.

Criterion 4: Observations with negative coordinates have clearly happened by mistake. Hence, these tracks have been reviewed.

Criterion 5: The resulted pedestrian trajectories have been checked and pedestrians with unexpected changes in movement direction or any kind of strange or unrealistic movements have been reviewed.

Normally, pedestrians do not walk in a straight line. Lateral oscillation (body sway) is a natural pattern in pedestrian's walking behaviour. This pattern changes according to density and the speeds of pedestrians. As the speeds of pedestrians decrease, the magnitude of oscillation increases, while its frequency decreases. Conversely, in low density situations when velocity is higher and there are not many interactions between pedestrians, the magnitude of this fluctuation remains constant and pedestrians can adjust their body pose more regularly (frequency of approximately 2 Hz) (Liu *et al.* 2009). The pedestrian trajectories are smoothed in order to lessen the effect of lateral oscillation on pedestrians' trajectories and also to decrease the errors resulted from camera calibration. A window size of 2 was selected to filter the data. That means the running averages of two successive positions of pedestrians are calculated to smooth the data.

By reviewing the data for the second time (in some cases third time), it is assured that the error resulted from manual tracking process has been reduced to the minimum and consequently extracted information regarding speed, direction, and the position of pedestrians is reliable. All data has been collected every half a second; data for 609 pedestrians was collected and a total of 11343 positions were observed.

4.2.5 Extraction of Choices

It is essential to process the observed positions and identify the choices that are available every time step along with choices that have been made. Subsequently, the choice set for each pedestrian is built by projecting positions for a time step of one second. Hence, all the variables

in the utility function have been measured and the selected choice which represents the location of the pedestrian after 1 second has been specified.

The destination of each pedestrian is her position in the last frame where she has been tracked. That is, while the actual destination of the pedestrian is unobservable in this video, the destination for utility evaluations was considered to be the last tracked position.

The data regarding the motion of pedestrians in groups has been collected. Group formation was recognized by observing the distances that pedestrians maintain from each other. Note that only those pedestrians that walk together during the whole observation period were labeled as groups. Observation of communication signs between these pedestrians was used as a confirmation criterion in this process. Groups were mostly consisted of a small number of people; yet, large groups were not observed. Out of a total of 191 observed groups, respectively 148, 32, 8, and 3 groups of size 2, 3, 4, and 5 have been observed. Generally, group members match their speed and direction with those of other nearby group members. However, in this data groups are mostly consisted of a small number of people. It is therefore assumed that all pedestrians maintain mutual awareness and therefore adapt their walking characteristics with all other people in the group. Since group members are only observed for a limited period of time, during which they are visible in the video sequence, without lack of accuracy it is assumed that they do not leave or change their groups during the observation period.

Note that not all observations could be used to calibrate and validate the model and therefore some of them have been removed from the data. Observation removal is performed whenever they meet any of the following criteria:

Criterion 1: The chosen location for the next time step is outside the defined choice set, *i.e.,* is not inside the boundaries of any of the alternatives. This contains observations related to stationary pedestrians having the speed of zero. This happens because any location of the pedestrian in the next time step is outside the choice set with current speed of zero (choice set of radius zero) or is exactly located on the current position (in case the pedestrian do not move in the next time step as well). Also, this may happen as a result of an unexpected change in the speed of pedestrians. That means when a pedestrian largely increases her speed from one time step to the next step, the location she occupies in the next time step exceeds the boundaries of the choice set.

Criterion 2: In this model, interactions between the decision-maker and other pedestrians are not restricted to local interactions. Instead, the decision-maker's behaviour is studied inside a region of interest in front of her. Therefore, some of the factors impacting the behaviour of those pedestrians walking close to the boundaries of the video frames cannot be observed and captured. For example, the leader-follower behaviour is investigated in different radial cones with the length of 5 times the radius of the choice set in front of each pedestrian. Therefore, when the pedestrian is near the boundaries of the video frame, the leaders inside the discussed area cannot be observed, *i.e.,* the pedestrian may have a leader which cannot be observed and consequently its effect cannot be captured in the model. Hence, a buffer zone has been considered, outside which the observed choices would be ignored in order to mitigate this problem. The defined buffer zone is illustrated in Figure 4.4. The black lines show the boundaries of the buffer zone. The direction of the prevalent flow is shown with the blue arrow in the figure and is from bottom towards the top of the video frame. Very few movements were observed in the opposite direction. Therefore, no buffer area was defined at the bottom of the

frame because most pedestrians pass the bottom boundaries of the frame to enter the area, not to exit it. Therefore, the potential leaders (or colliders) for each observable pedestrian in the frame are in front of her, inside the displayed buffer zone, and those pedestrians close to the bottom boundaries of the frame can only be potential leaders for pedestrians that are not still observable in the frame (their choice behaviours are not studied yet). The impact of the leader (or collider) on walking behaviour of pedestrian decreases with the distance between them. Therefore, considering this buffer zone, even if a pedestrian has a leader (or collider) that cannot be observed, it can be assured that her effect on pedestrian's decision making would be very low due to the large distance between her and the pedestrian.

Criterion 3: When one or more members of a specific group are not visible in a frame, the choices made by present members have been excluded. This is because the influence of absent members on present ones cannot be captured in these cases.

It should be noted that even though in all the above situations the choice selected by the decision-maker is not used to estimate or validate the model, the pedestrian's information in that specific time step can still be used to model other pedestrians' behaviour. That is, the effect of the pedestrian on *modelling* walking behaviour of other pedestrians is still considered. For example, the pedestrian can be a leader for another pedestrian nearby. After eliminating those observations belonging to above criteria, 8343 choice observations were used to calibrate and validate the model.

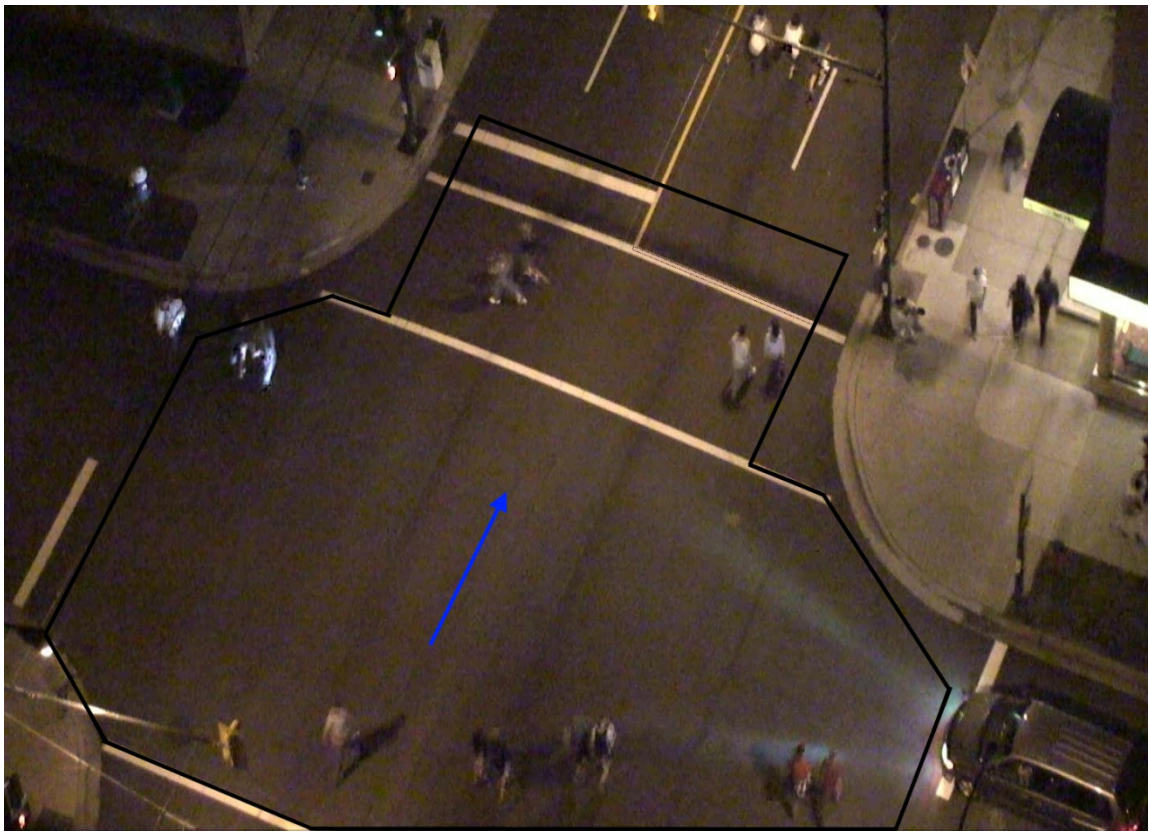


Figure 4.4: The buffer zone

4.3 Model Calibration and Validation

In this section the results are reported and discussed. First, some data statistics are presented in sub-section 4.4.1, then the estimation results are reported and discussed in sub-section 4.4.2, and finally in sub-section 4.4.3 the validation process is described and results are presented.

4.3.1 Data Statistics

Summary of the resulted data obtained from video recordings is reported in this sub-section. Two examples of manually tracked pedestrian trajectories are shown in Figure 4.5.

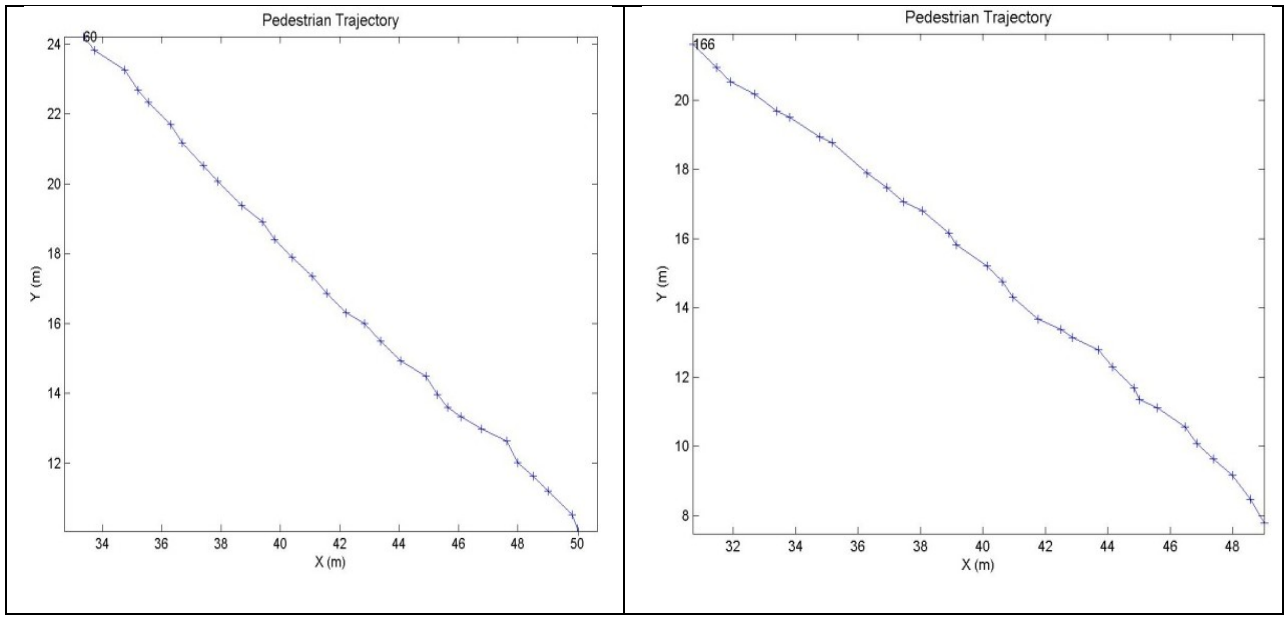


Figure 4.5: Examples of pedestrian trajectories (horizontal and vertical axes are x and y world coordinates respectively)

Table 4.1 shows some sample data statistics for 20 different pedestrians. The average speed, standard deviation of speed, total walking distance for pedestrians belonging to groups and individual pedestrians are reported. The indicator I_g, as described in Equation 4.1, is 1 if pedestrian is a group member and zero otherwise. Histogram of observed speeds for all pedestrians in all time steps is illustrated in Figure 4.6. Some of the speed statistics are reported in Table 4.2. All speed values are stated in *m/s*. Maximum observed speed is 3.74 *m/s*.

The number of pedestrians walking in groups is 439 pedestrians out of a total of 609. Only 28% of pedestrians walk individually. The speed data for people walking in groups (9302 observations) and those walking individually (3254 observations) are reported in Table 4.3 and Figure 4.7. The maximum speed of individual pedestrians is higher than the maximum speed of pedestrians walking in group. There is no apparent difference between mode and average speed for both categories. The difference between the standard deviation of speeds of individual

pedestrians and pedestrians walking in groups is statistically significant (F-statistic is 1.47 and p-value is less than 0.001), however it appears only to be slightly higher for individual pedestrians.

Table 4.1: Sample data statistics

Pedestrian	I_g (1 if Group member)	Walking Distance (m)	Average Speed (m/s)	Speed Standard Deviation (m/s)
1	1	19.1535	1.6596	0.2208
2	1	21.46	1.3199	0.3492
3	1	22.1788	1.3744	0.1599
4	1	25.1289	1.5886	0.2246
5	1	25.6125	1.617	0.2551
6	0	15.2883	1.6638	0.0964
7	0	16.1447	1.4039	0.2668
8	1	15.1675	1.1684	0.1959
9	1	18.6348	1.7088	0.1527
10	1	21.0002	1.2727	0.1612
11	0	14.7662	1.3424	0.2991
12	1	21.8584	1.327	0.2282
13	1	22.1134	1.3429	0.2232
14	1	21.2638	1.3373	0.23
15	1	22.1736	1.3596	0.2636
16	0	16.2846	1.3028	0.2383
17	0	22.4171	2.1351	0.4887
18	1	22.2804	1.3665	0.1946
19	1	21.46	1.3199	0.3492
20	1	22.1788	1.3744	0.1599

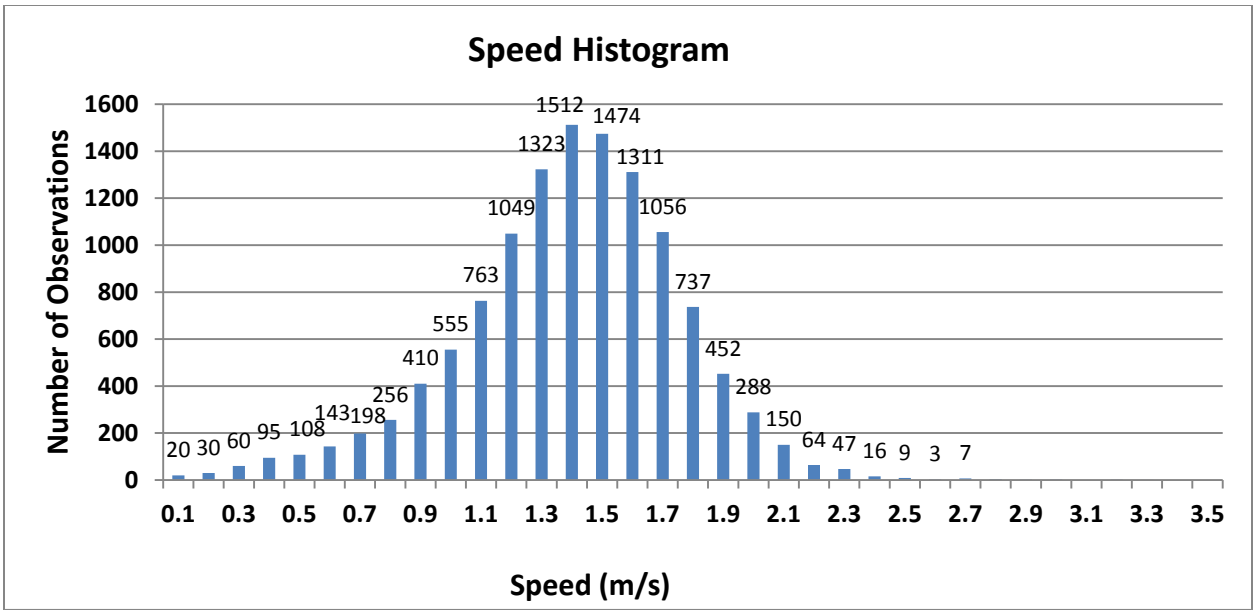


Figure 4.6: Speed Histogram

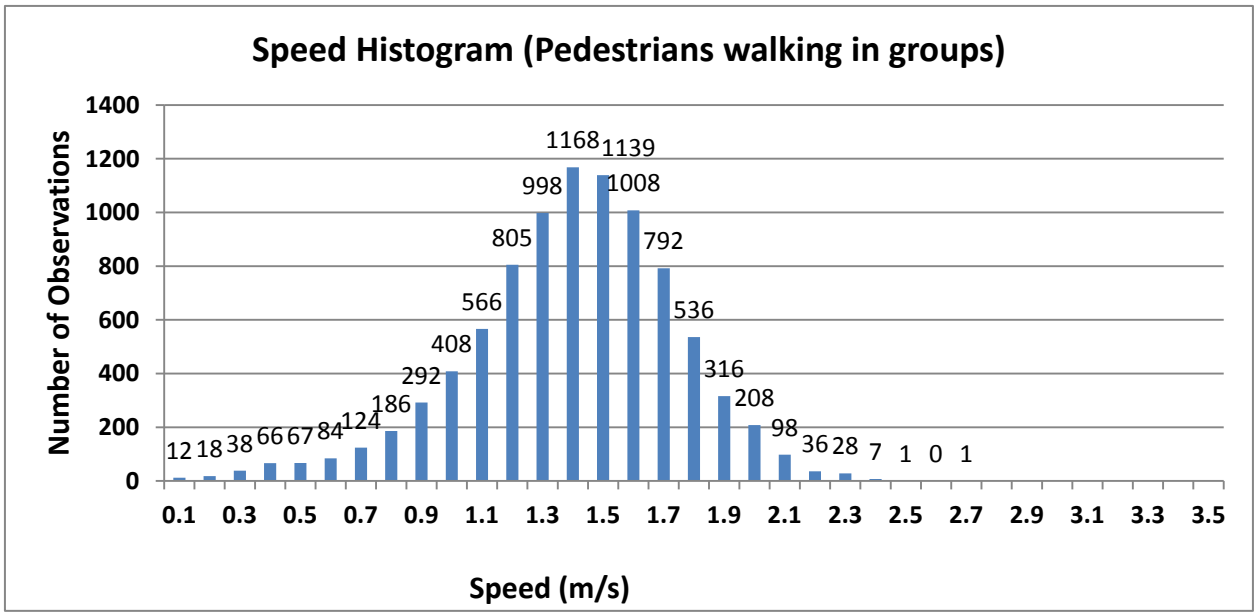
Table 4.2: Speed Statistics

Max observed speed (m/s)	3.234
Min observed speed (m/s)	0
Average speed (m/s)	1.302
Standard Deviation (m/s)	0.433
Most frequent observed speed (m/s)	1.4
Median of speed data (m/s)	1.358

Table 4.3: Speed data for group member and individual pedestrians

	Pedestrians in group	**Individual pedestrians**
Max observed speed (m/s)	2.6176	3.2341
Min observed speed (m/s)	0	0
Average speed (m/s)	1.347	1.344
Standard Deviation (m/s)	0.347	0.421
Mode (m/s)	1.4	1.3
Median (m/s)	1.373	1.367

a)



b)

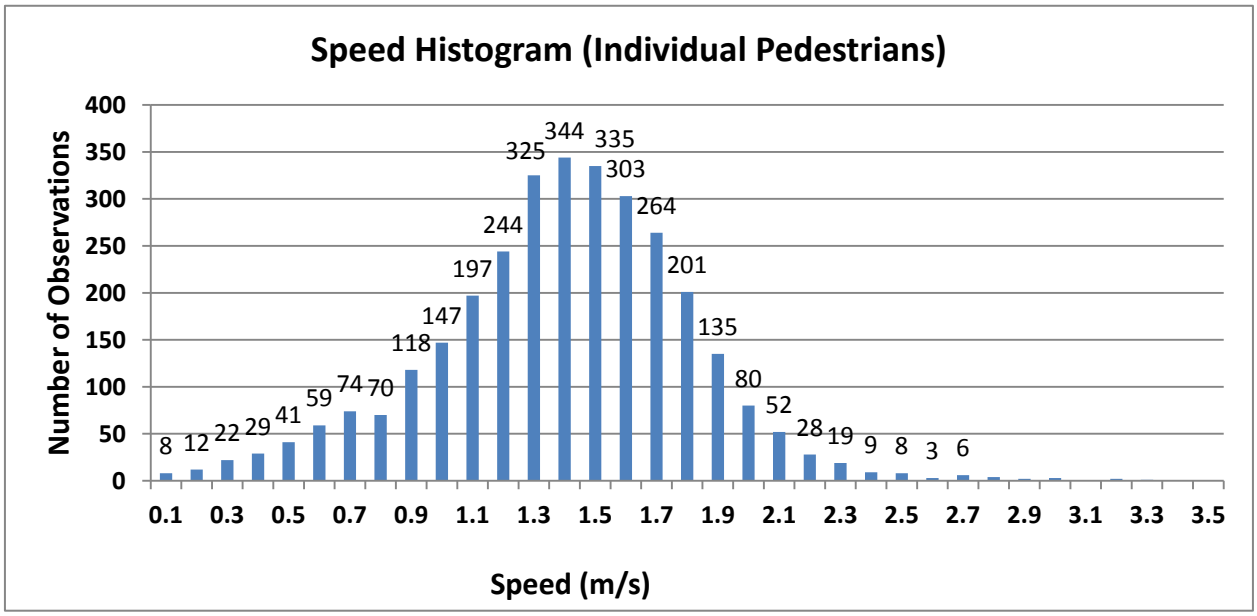


Figure 4.7: Speed Histogram of (a) pedestrians in groups (9302 observations) and (b) pedestrians walking individually (3254 observations)

Figure 4.8 illustrates the change in the movement direction of pedestrian trajectories based on 3 angle ranges representing small, moderate, or high change in direction. The angle ranges are defined as follows:

$$\begin{cases} 60° <= \text{Dir} <= 120° & \text{no / small direction change} \\ 0° <= \text{Dir} < 30° \quad \textbf{or} \ 150° < \text{Dir} <= 180° & \text{moderate direction change} \\ 30° <= \text{Dir} < 60° \ \textbf{or} \ 120° < \text{Dir} <= 150° & \text{high direction change} \end{cases}$$

It can be seen that pedestrians tend to keep their direction and only few cases of high or moderate change in movement direction have been observed (4.44%, 6.84%, and 88.72% of high, moderate, and small direction change respectively).

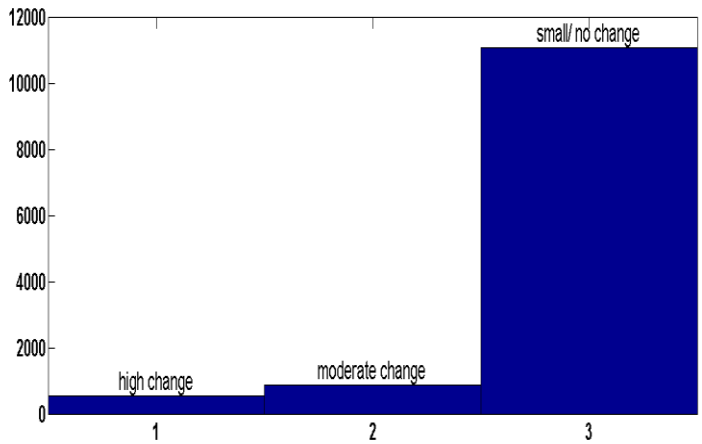


Figure 4.8: Histogram of different extents of change in movement direction

The same approach has been performed for group people and individual pedestrians as well and similar results have been obtained in both cases. For individuals, high, moderate, and small direction change of 5.10%, 7.55%, and 87.35% and for pedestrians walking in groups, high, moderate, and small direction change of 4.21%, 6.60%, and 89.19% have been observed respectively. In order to figure out whether the proportions of different extent of change in direction are significantly different for the two categories (individuals and group members) the

chi-square test has been performed. The results of this statistical test imply that the proportions of small, moderate, and high direction change for group members are significantly different from those of individuals at the 0.5%, 10%, and 5% level of significance respectively (p-values of 0.005, 0.072, and 0.039 respectively).

The histogram of observed choices is illustrated in Figure 4.9. Expectedly, the choice associated with cell 17, which corresponds to the selection of constant velocity, is the most frequent selected choice. Choices situated at the extreme side of the choice set are rarely chosen. Table 4.4 illustrates observed selected choices in each nest. Expectedly, alternatives belonging to the nest with constant speed are those that have been selected the most often.

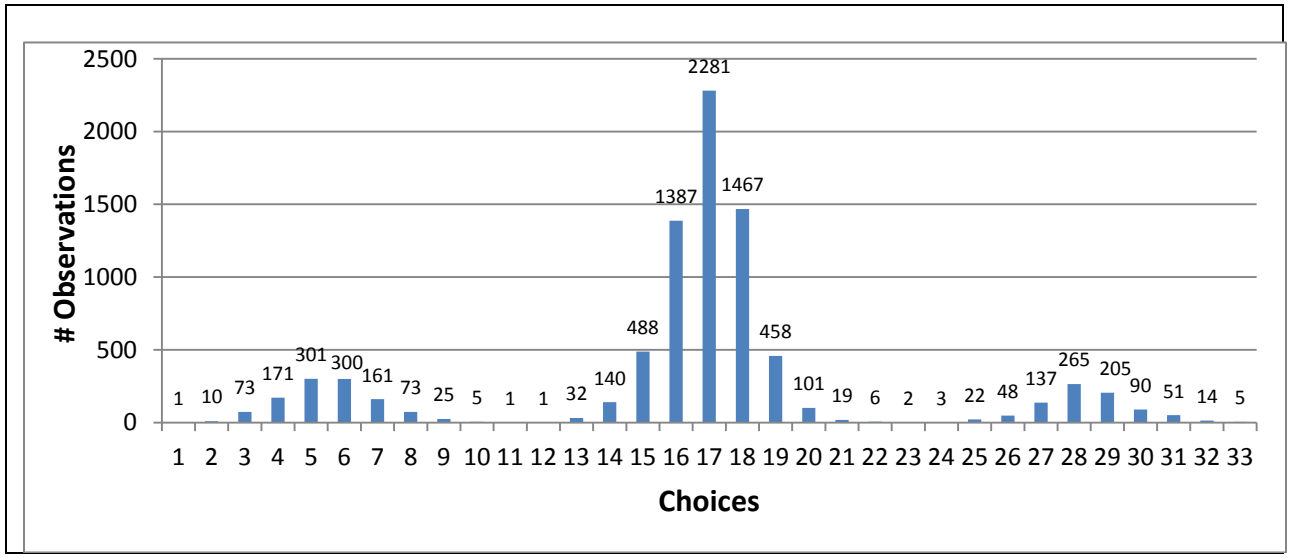


Figure 4.9: Histogram of observed choices

Table 4.4: Number of selected choices in each nest

Nests	Number of Observations	Percentage of Total
Acceleration	1121	13.43641376
Deceleration	842	10.09229294
Constant Speed	6380	76.4712933
Central	2846	34.11242958
Non-central	5497	65.88757042

From 8343 observation used to calibrate the model, accelerated leaders, decelerated leaders, and colliders are present in 4062 (48.69%), 5044 (60.47%), and 5370 (64.36%) of them respectively. Therefore, a majority of the leaders in this dataset are decelerated leaders whose speeds are less than decision-makers' speed. The number of observations in which leaders or colliders are available for each of the 11 radial cones is reported in Table 4.5.

A total of 6254 observations out of these 8343 observations (approximately 75%) are associated with group members and the remainder belong to individual pedestrians. With respect to acceleration behaviour, the portion of observations for pedestrians with the speed less than 1.39 m/s is almost the same size as the portion with the speed more than 1.39 (49% vs. 51%).

Table 4.5: Presence of leaders and colliders in each radial cone

Radial Direction	Accelerated Leader		Decelerated Leader		Collider	
	Number of observations	Percentage of total	Number of observations	Percentage of total	Number of observations	Percentage of total
1	200	2.397	464	5.562	1568	18.794
2	128	1.534	671	8.043	1432	17.164
3	320	3.836	780	9.349	1031	12.358
4	779	9.337	891	10.680	788	9.445
5	1310	15.702	1535	18.399	690	8.270
6	1499	17.967	1904	22.822	642	7.695
7	1124	13.472	1606	19.250	654	7.839
8	501	6.005	954	11.435	680	8.151
9	164	1.966	543	6.508	899	10.776
10	100	1.199	234	2.805	1142	13.688
11	222	2.661	276	3.308	1305	15.642

4.3.2 Model Calibration

In this sub-section the estimation results are reported. For the sake of comparison, two separate models have been estimated, one with group behaviour and one without group behaviour. The utility parameters are computed using the trajectory data extracted from video recordings and are entered as inputs to the estimation tool. Models are estimated by applying maximum likelihood estimation using Biogeme (Bierlaire 2003). Among optimization algorithms available in Biogeme, the 'DONLP2' was used in this study. All available algorithms have their own advantages and disadvantages and may differ in terms of speed of convergence. DONLP2 was appropriate for our purpose with a reasonable convergence time. This algorithm was developed by Spellucci (1993) which is a sequential equality constrained quadratic programming method. The description about this algorithm can be found in Spellucci (1998a) and Spellucci (1998b). A total of 8343 choice observations are used to calibrate both models. The values of estimated parameters are shown in Table 4.6 and Table 4.7 for Model 1 and Model 2, respectively. Model 1 refers to the model without group behaviour and Model 2 refers to the model with group behaviour. The values of final log likelihood at convergence for both models are reported in Table 4.8.

Table 4.6: Estimated parameters for Model 1 (without group behaviour)

Variable Name	Value	t-test	p-value	Variable Name	Value	t-test	p-value
β_{accHS}	-38.2	-2.12	0.03	δ^L_{acc}	0	--fixed--	
β_{accLS}	-3.12	-21.37	0.00	δ^L_{dec}	-0.0731	-4.94	0.00
β_{ddir}	-0.0726	-19.92	0.00	γ^L_{acc}	0.319	2.27	0.02
β_{ddist}	-1.83	-11.71	0.00	γ^L_{dec}	1.02	9.17	0.00
β_{dec}	-0.189	-4.42	0.00	ρ^L_{acc}	0	--fixed--	
$\beta_{dir\_central}$	-0.00665	-2.76	0.01	ρ^L_{dec}	-0.261	-1.87	0.06
$\beta_{dir\_extreme}$	-0.0252	-7.17	0.00	λ_{accHS}	1.88	4.56	0.00
β_{dir_side}	-0.0143	-8.73	0.00	λ_{accLS}	2.99	14.51	0.00
α_C	-0.00174	-2.56	0.01	λdec	-0.963	-8.45	0.00
ρ_c	-0.0169	0.4	0.69	$\mu_{acceleration}$	1.17	2.59	0.01
δ_c	1	--fixed--		$\mu_{central}$	3.26	1.52	0.13
γ_c	0	--fixed--		$\mu_{constant}$	1.3	5.24	0.00
α^L_{acc}	0.52	5.17	0.00	$\mu_{not\_central}$	2.6	8.45	0.00
α^L_{dec}	1.95	4.45	0.00	$\mu_{deceleration}$	1	--fixed--	

Table 4.7: Estimated parameters for Model 2 (with group behaviour)

Variable Name	Value	t-test	p-value	Variable Name	Value	t-test	p-value
β_{accHS}	-35.9	-2.07	0.04	γ^L_{acc}	0.357	2.41	0.02
β_{accLS}	-3.09	-20.65	0.00	γ^L_{dec}	1.11	8.98	0.00
β_{ddir}	-0.0729	-21.15	0.00	ρ^L_{acc}	0	--fixed--	
β_{ddist}	-1.72	-11.93	0.00	ρ^L_{dec}	-0.29	-1.85	0.06
β_{dec}	-0.18	-4.49	0.00	λ_{accHS}	1.87	4.42	0.00
$\beta_{dir\_central}$	-0.0041	-1.58	0.11	λ_{accLS}	3.07	14.09	0.00
$\beta_{dir\_extreme}$	-0.0192	-5.61	0.00	λ_{dec}	-0.953	-8.59	0.00
β_{dir_side}	-0.0105	-6.21	0.00	β_{gdir}	-0.027	-3.75	0.00
α_C	-0.0017	-2.45	0.01	β_{gdist}	-0.491	-9.48	0.00
ρ_c	-0.0136	0.32	0.75	δ_g	0.95	12.76	0.00
δ_c	1	--fixed--		ρ_g	1.19	16.64	0.00
γ_c	0	--fixed--		$\mu_{acceleration}$	1.11	1.87	0.06
α^L_{acc}	0.532	5.12	0.00	$\mu_{central}$	2.3	2.03	0.04
α^L_{dec}	1.9	4.02	0.00	$\mu_{constant}$	1.22	4.07	0.00
δ^L_{acc}	0	--fixed--		$\mu_{not\_central}$	2.36	7.53	0.00
δ^L_{dec}	-0.075	-4.12	0.00	$\mu_{deceleration}$	1	--fixed--	

Table 4.8: Maximum Likelihood Estimation summary

Model 1 : Without Group Behaviour	**Model 2 : With Group Behaviour**
Number of observations: 8343 Number of estimated parameters : 23 Final log-likelihood : -15481.273 Final gradient norm: 2.430e-006	Number of observations: 8343 Number of estimated parameters : 27 Final log-likelihood : -15266.451 Final gradient norm: 6.017e-008

As can be seen in the previous table, adding more parameters to the model resulted in a higher likelihood score. In order to find out if adding the newly presented parameters significantly improves the fit of the model to the dataset, Likelihood ratio test has been performed. It is required to examine whether the improvement in Log-likelihood of 214.822 is sufficient to justify the inclusion of 4 more parameters to the model. The test statistic, Likelihood ratio (LR), is computed as follows:

LR= 2*(log-likelihood Model 1-log-likelihood Model 2) = 2* (15481.273-15266.451) =429.644

Degree of freedom is the number of added parameters to Model 2 and is equal to 4 (27-23). With this information and knowing that LR almost follows the Chi-square distribution, the critical value at p=0.001 (99.9 % level of confidence) is obtained from the chi-square tables and is equal to 18.467. The difference between the likelihood estimation of two models (429.64) is much higher than this critical value, which implies that Model 2 (developed model with group behaviour) fits the dataset significantly better. In other words, it can be concluded that incorporating the effect of grouping relationships into the walking behaviour model significantly improves its fit to the data. Therefore Model 2 is the model that is selected to explain the data in this research. This signifies an important finding in this thesis.

The values and signs of estimated parameters in Table 4.7 are consistent with what was expected. Parameters β_{ddist} and β_{ddir}have negative signs implying that the attractiveness of an alternative decreases as the distance or direction to destination gets larger. This shows the tendency of the pedestrian to select the shortest path to reach to the destination. Parameters $\beta_{dir\_central}$, β_{dir_side}, and $\beta_{dir\_extreme}$ have negative signs indicating that the decision-maker prefers to keep the current direction and if possible tends to avoid frequent direction changes. The value of $\beta_{dir\_central}$ is higher than β_{dir_side} and the value of β_{dir_side} is higher than $\beta_{dir\_extreme}$. This implies that regarding this behaviour, the attractiveness of alternatives at the centre of the choice set (small/no direction change) is higher than that of alternatives with moderate direction change (side cones) which in turn is higher than the attractiveness of alternatives with high direction change (extreme cones). Parameters β_{dir_side} and $\beta_{dir\_extreme}$ are significant; $\beta_{dir\_central}$ is also significant but at a lower level of confidence (89%), which can suggest the removal of this parameter from the utility function (fix to zero) so that the central direction can be used as a reference direction, fluctuation from which leads to the penalty of reducing the utility value. Parameters β_{dec}, β_{accLS}, and β_{accHS} have negative values indicating decision-maker's willingness to walk at a constant speed (free flow acceleration term for alternatives with constant speed is equal to zero). Parameters λ_{accHS} and λ_{accLS} are positive, pointing out the fact that at higher speeds the tendency of pedestrians to accelerate is less compared to lower speeds and conversely, λ_{dec} is negative, showing that when the speed increases the tendency of pedestrians to decelerate increases. The absolute value of β_{accHS} is higher than two other parameters β_{dec} and β_{accLS}. Parameters v_{max} and v_{maxLS} are respectively fixed to 4.84 m/s and 1.39 m/s (Robin *et al.* 2009).

Parameters α_{acc}^{L} and α_{dec}^{L} have positive signs which capture the attractiveness of the alternative for which the leader has been identified relative to the alternative with no leader; having a leader motivates the decision-maker to accelerate or decelerate depending on the leader's speed. It is assumed that the presence of a leader has no influence on the utility of alternatives with constant speed. Parameter ρ_{dec}^{L} is negative showing that the further the leader is from the decision-maker the less the impact of the leader on decision-maker's choice and deceleration behaviour would be. Parameter ρ_{acc}^{L} is not statistically significant and therefore is set to zero. Parameters γ_{acc}^{L} and γ_{dec}^{L} have positive values which means the greater the difference between the leader and decision-maker's speed is, the greater the tendency of the decision-maker to accelerate or decelerate will be and consequently the utility of correspondent alternatives increases. Parameter δ_{dec}^{L} has a negative sign expressing the fact that as the difference between the movement direction of leader and the direction of the alternative increases, the effect of the leader on decision-maker's behaviour decreases. Parameter δ_{acc}^{L} is not significantly different from zero and is fixed to zero. All leader-follower parameters related to deceleration are significant. On the other hand, associated parameters to relative distance and the direction of accelerated leaders are not significant. One potential reason for this is that the leader-follower behaviour of pedestrians in this dataset is better explained through group maintenance than following pedestrians ahead in space. Another reason may be that the larger portion of leaders identified in this dataset are decelerated leaders rather than accelerated ones. Hence, having a larger dataset with enough information regarding accelerated leaders may overcome this problem.

Parameter α_c has a negative sign indicating that the attractiveness (utility) of the alternative decreases if a collider is identified for the cone in which the alternative locates. Parameters γ_c and δ_c are not statistically significant and are fixed to 0 and 1 respectively. Parameter ρ_c is not

significantly different from zero and can be removed from the model. Its negative sign implies that as the distance of the collider and the centre of the alternative gets bigger, the effect of the collider on the utility of the alternative gets smaller. The reason that 3 out of 4 parameters related to collision avoidance behaviour are not statistically significant is the lack of information related to this behaviour in the data. This dataset consists of the unidirectional flow in which most pedestrians walk in the same direction rather than walking in opposite directions (which is the direction movement of a potential collider in this model).

All 4 estimated parameters for group behaviour are significantly different from zero (shown in blue color in Table 4.7). Parameters β_{gdist} and ρ_g have negative and positive signs respectively, reflecting the fact that the more the distance between the centre of the alternative and the group centre is, the less the attractiveness and the utility of the alternative would be. Similarly, β_{gdir} and δ_g have negative and positive signs respectively showing that the closer the direction of the alternative to the average moving direction of group members is, the more attractive the alternative would be. Except for ρ_c and the four fixed parameters described earlier, the rest of estimated parameters are significant as shown in Table 4.7.

As mentioned before, the error structure captures the uncertainty and unobserved factors in the modelling process of decision making. The error structure in this study is captured by CNL model to take into account the correlation pattern between alternatives within the choice set. Five nests have been defined and each alternative belongs to two nests. Therefore, the unobserved components of alternatives in different nests can be correlated and are captured by CNL model. Nest parameters indicate the degree of correlation between alternatives in each nest. The nest parameters of 4 out of 5 nests are significantly different from one and deceleration nest

parameter is fixed to one. Variations of different utility terms with respect to associated parameters are shown in Figure 4.11 to Figure 4.14.

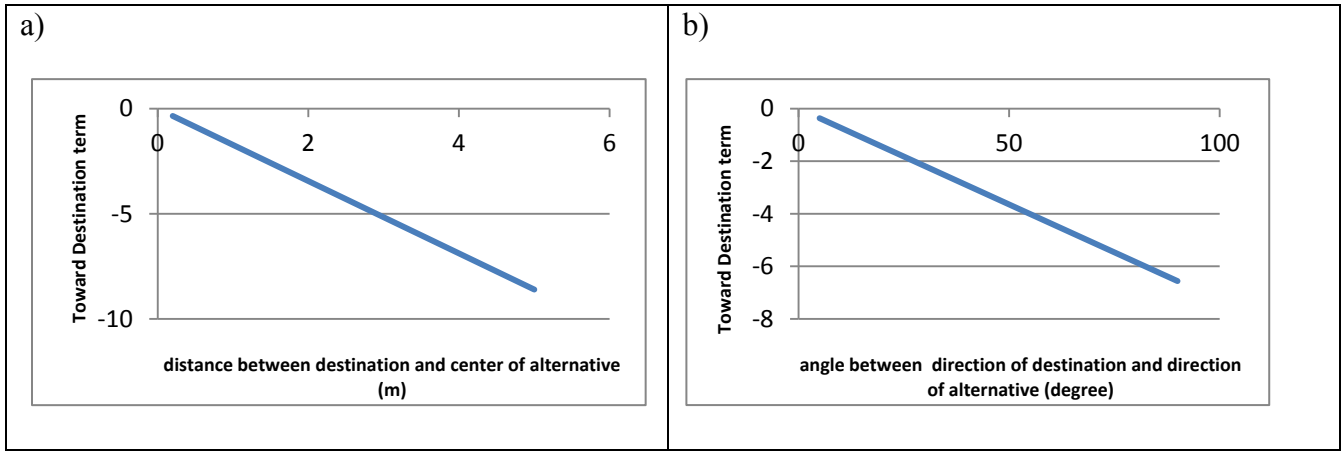


Figure 4.10: Effect of toward destination behaviour on utility function

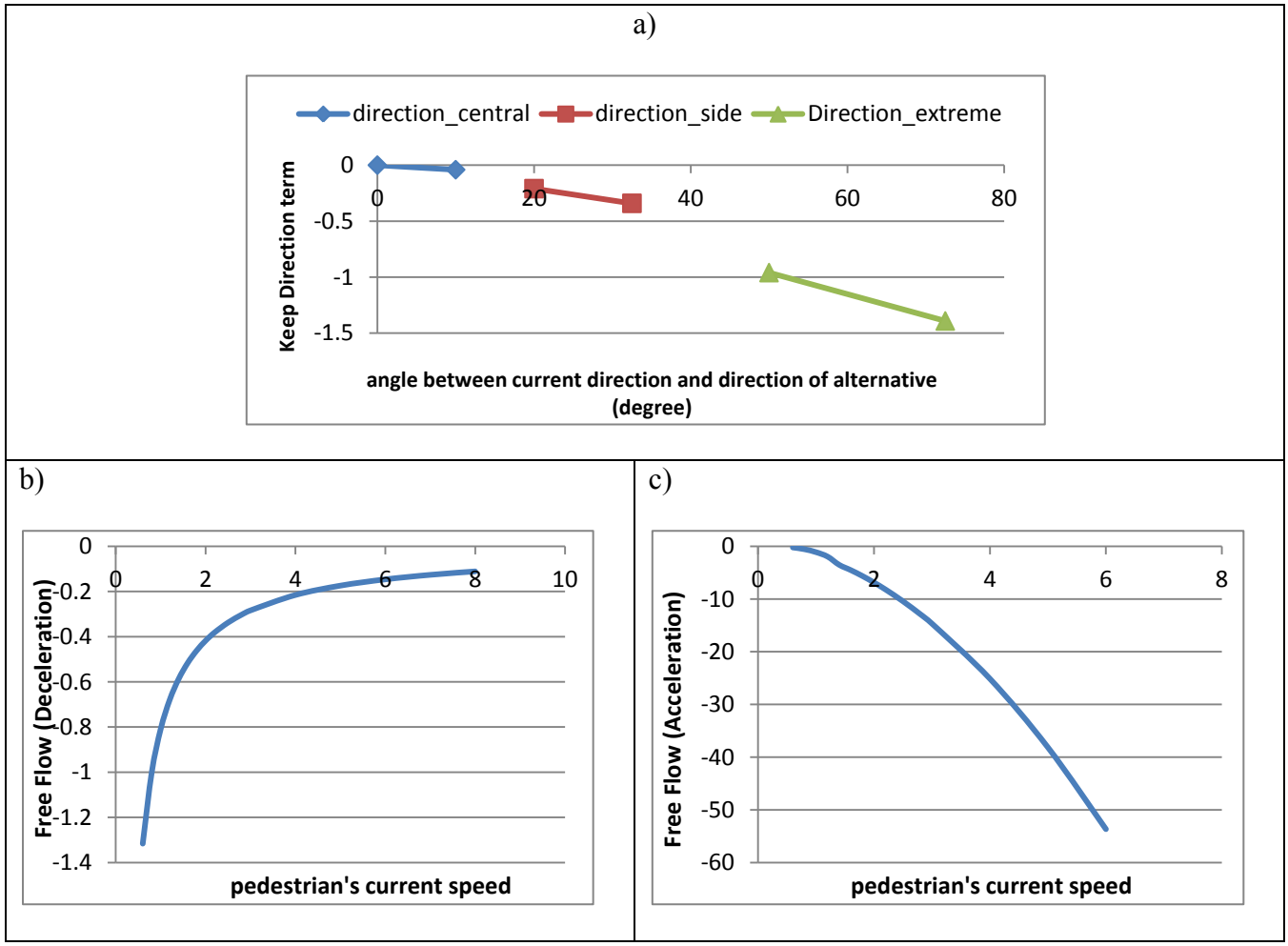


Figure 4.11: Effect of a) keep direction behaviour, b) free flow behaviour (Deceleration), c) free flow behaviour (Acceleration) on utility function

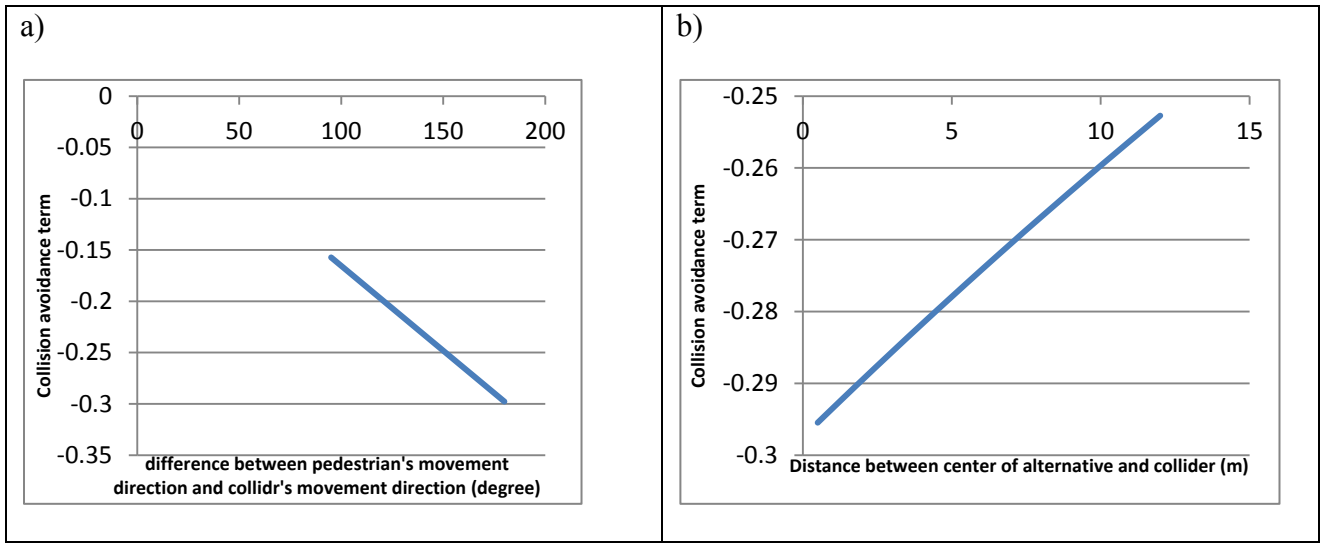


Figure 4.12: Effect of collision avoidance behaviour on utility function

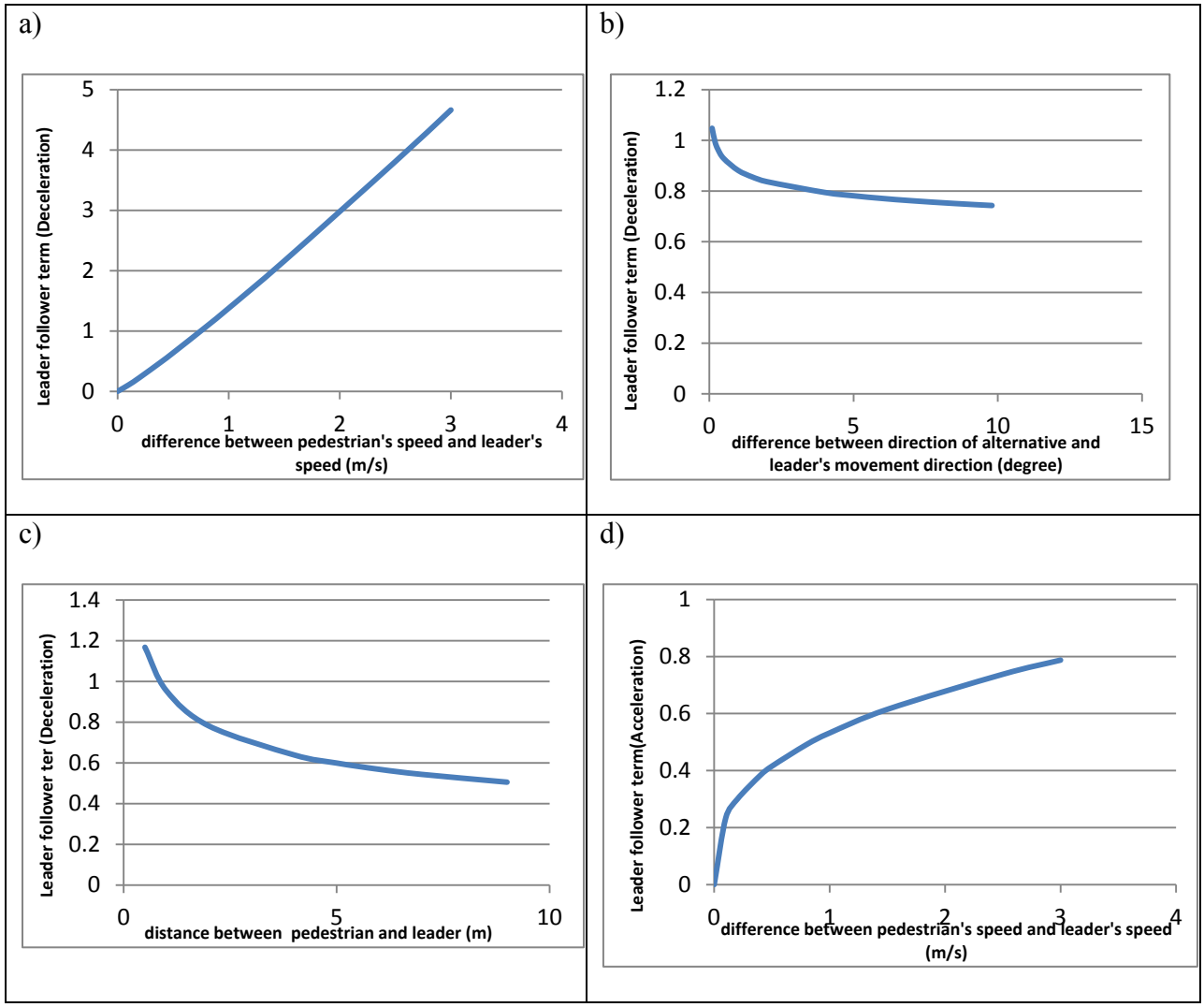


Figure 4.13: Effect of leader-follower behaviour on utility function

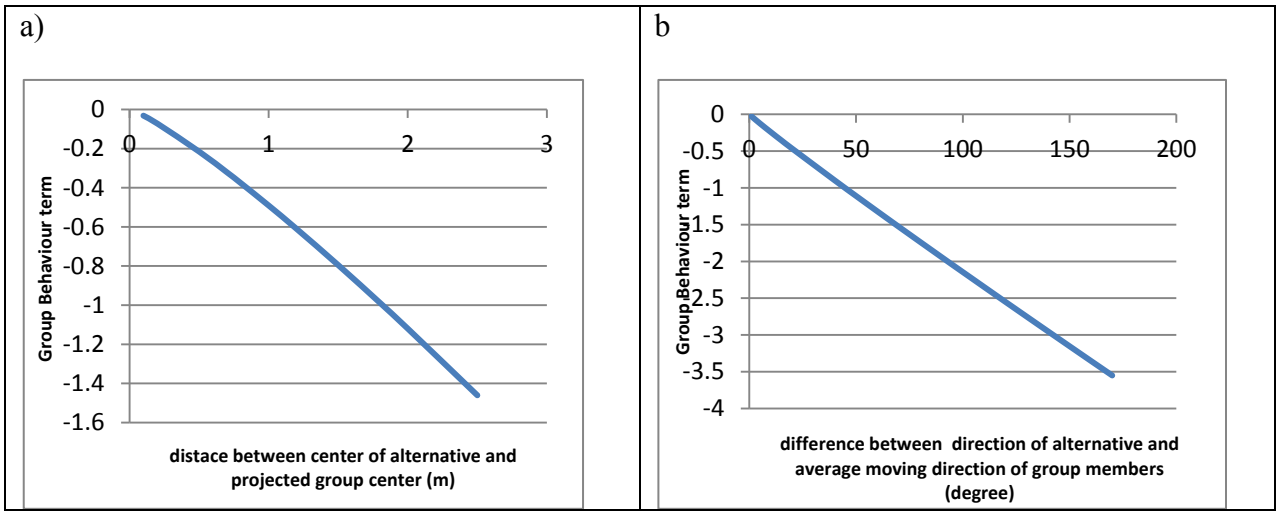


Figure 4.14: Effect of group behaviour on utility function

4.3.3 Validation

It is selected to adopt a consistent validation methodology with Robin *et al.* (2009) in order to enable cross-comparison of models and establish the findings in this thesis. Biosim (Bierlaire 2003) is used to simulate the model with estimated parameters (reported in Table 4.7). In order to compare the real observed choices to predicted choices, the model is performed on the whole dataset. From the results prepared by Biosim, for observation n the probability of selecting alternative i, given the estimated parameters, is obtained and is denoted as P_{in}. Therefore, the number of occurrence of each alternative can be obtained by adding all values of P_{in} over all observations ($\sum_n P_{in}$). The histogram of predicted choices by the model is shown in Figure 4.15 and the histogram of real observed choices is shown in Figure 4.9. The values are also shown in the figures for the sake of comparison.

It can be seen that there is no major difference between two histograms. The choice set is divided into different groups of alternatives regarding speed and direction and the numbers of predicted and actual observed choices in each group is compared. Results are reported in Table 4.9.

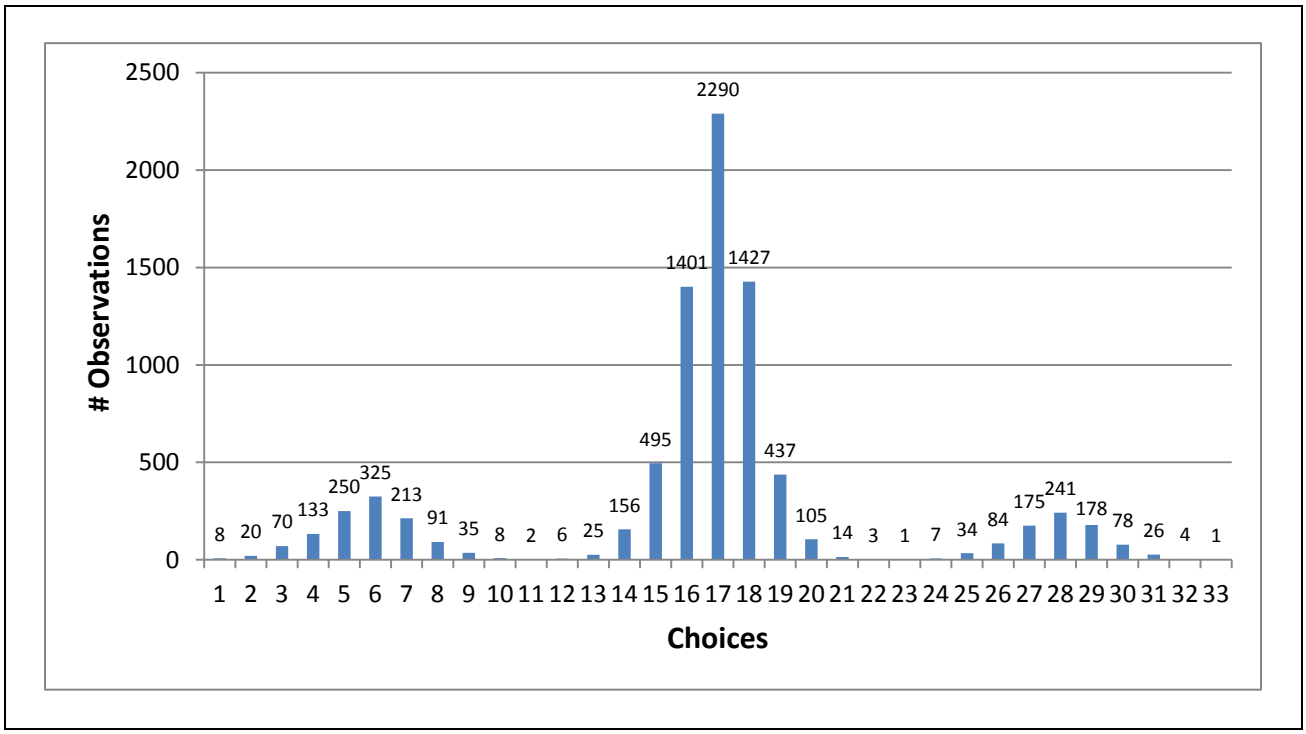


Figure 4.15: Histogram of predicted choices by the model

Table 4.9: Number of actual and predicted observations for groups of alternatives

Group	Alternatives	Predicted # of observed choices	Actual # of observed choices	Relative Error
Acceleration	1-11	1154.074	1121	0.029504024
Constant Speed	12-22	6360.645	6381	-0.003189918
Deceleration	23-33	828.281	842	-0.01629352
Front	5,6,7,16,17,18,27,28,29	6500.031	6504	-0.000610247
Left	3,4,14,15,25,26	971.925	942	0.03176769
Right	8,9,19,20,30,31	773.299	798	-0.030953812
Extreme left	1,2,12,13,23,24	66.292	49	0.352918111
Extreme Right	10,11,21,22,32,33	31.452	50	-0.370959253

Relative error is equal to the difference of predicted and actual observed choices divided by the number of actual observed choices. As it can be seen, errors are acceptable. The higher relative

errors in extreme left and extreme right groups are due to the small number of observations in these groups.

The histogram of the probabilities of selected choices predicted by the model is shown in Figure 4.16. Outliers are considered to be those observations with the probability of selected choice being less than 1/33, where 33 is the number of alternatives. This value represents the probability of selecting an alternative by mere chance where all the alternatives have equal probabilities to be chosen. The blue line in the Figure indicates this threshold and observations below the line are considered to be outliers. Outliers constitute 10.49% of the observations. With looking at the shape of the histogram, it can be said that the frequency of selected choices with higher probabilities are relatively higher than the frequency of selected choices with low probabilities (except for the outliers and the initial part of the histogram). This is an indication of the good behaviour of the model.

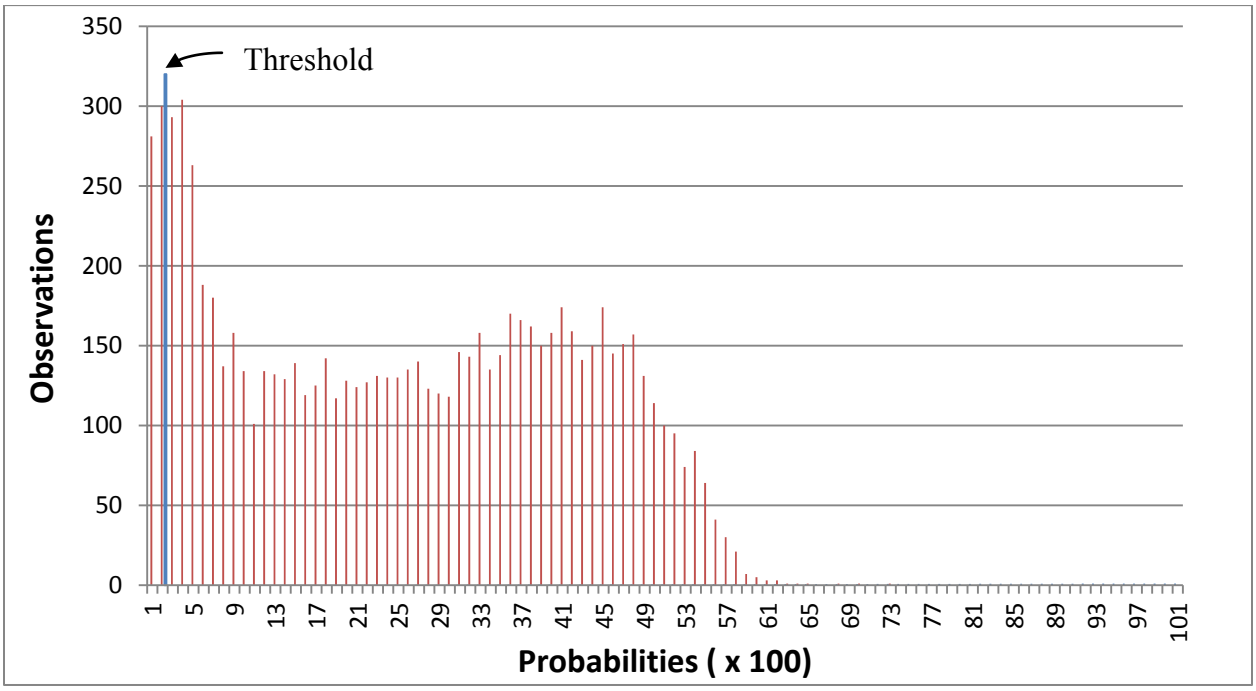


Figure 4.16: Histogram of probabilities of selected choices predicted by the suggested model

In order to examine the performance of the suggested model versus the constant only model, the number of outliers and log likelihood of the two models are compared. The constant only model is a model in which the utility of each alternative is presented just by an alternative specific constant (ASC). There are a few number of outlier observations in the suggested model (875 out of 8343). 16.73 % of observations (1396 out of 8343) are outliers in constant only model, which is more than the number of outliers in the suggested model (10.49%). The histogram of the probabilities of selected choices for ASC model is illustrated in Figure 4.17.

final log likelihood of the constant only model is equal to -19782.556 which is significantly less than log likelihood of the developed model (-15266.451). In total, the results obtained from comparing the developed model and constant only model in addition to the results obtained from comparing the actual and predicted observed choices indicate the good performance and acceptable specification of the suggested model and shows the better predictive power of the suggested model compared to ASC model.

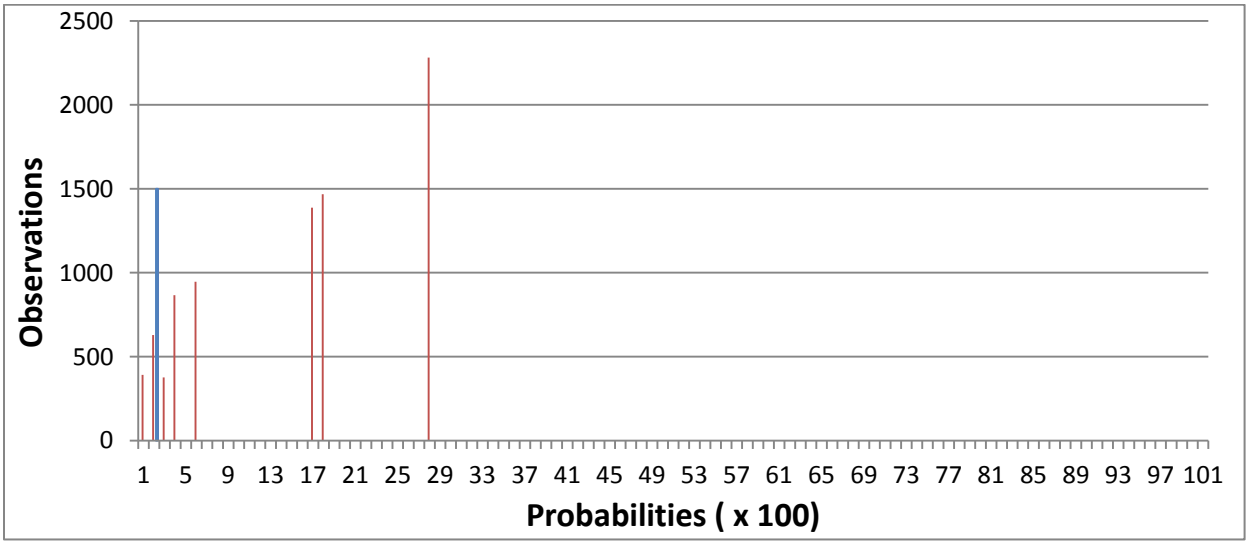


Figure 4.17: Histogram of probabilities of selected choices (ASC model)

Another approach to examine the model performance is to conduct cross-validation on the dataset. For this purpose, the observations are separated into 4 parts, each including 25% of the data. Four trials have been carried out and in each trial the model is calibrated on 3 parts of the dataset (75%) and then is validated on one remaining part. The numbers of outliers are measured for each trial. The same experiment has been performed with ASC model. The percentages of outliers for the suggested model for trial 1 to trial 4 are respectively: 9.53%, 10.15%, 10.44%, and 10.58% and for constant-only model are: 14.42%, 16.76%, 17.58%, and 18.15%.

It can be seen that for all the trials there are relatively few number of outliers in the suggested model which is less than the number of outliers in ASC model. Moreover, the percentage of outliers obtained from cross-validation is compatible with what is obtained from the validation of the full dataset. All of these results are an evidence of a stable model with reliable structure and specification.

4.4 Summary

The maintenance of group formation affects the walking behaviour of pedestrians. The decisions of pedestrians in groups are influenced by the behaviour of other group members. In this chapter, the impact of group behaviour on pedestrian's decision making was modeled using the discrete choice framework. It was found out that group behaviour plays an important role in building a walking behaviour model for a human crowd containing of groups of pedestrians. The correlation structure between alternatives has been captured with CNL specification. The estimation results show that the parameters associated with this behaviour are statistically significant and their values and signs are consistent with what was expected for this behaviour. The good performance and reliability of the model has been confirmed in the validation process.

Chapter Five: Investigation of Pedestrian Behaviour Passing through Bottlenecks in High Density Situations

Next step model was estimated based on bi-directional pedestrian flow and was validated regarding choices made by decision-makers. In this chapter, the behaviour of pedestrians when passing through bottlenecks in high density situations is investigated using DCM. For this purpose, first next step model is implemented in a simulation framework and later the required revisions are applied to the model to make it consistent with the mentioned situation. Finally, it is aimed to calibrate and validate the model based on real-world pedestrian data. Section 5.1 explains the simulation tool. In section 5.2, the real pedestrian data which is used for calibration and validation purposes is described. In Section 5.3, pedestrian walking behaviour at bottlenecks and in high density situations is simulated using next step model and the results are investigated. Sections 5.4 and 5.5 explain model developments and revisions respectively. The criteria employed to keep the simulation on track are described in section 5.6. The calibration and validation process and associated results are reported in sections 5.7 and 5.8 respectively. Finally, this chapter will end with a summary in section 5.9.

5.1 Simulation Tool

A simulation tool is developed to examine the performance of the discrete choice walking behaviour model in different situations and to observe whether the model represents the realistic walking behaviour or not. The simulation tool is written in Matlab. Positions and information associated with all pedestrians are updated every time step (time-based simulation) and the time step is selected to be half a second in this study.

The inputs of simulation tool include initial position, direction, and speed of each pedestrian; destination of every pedestrian; the time that each pedestrian enters the simulation; coefficients and parameters of the model; and the configuration of the walking area. Each pedestrian's trajectory, histogram of choices, and interpersonal distances are some of the most useful outputs of the simulation.

Each pedestrian has a final destination and may also be associated with a few intermediate destinations, depending on the configuration of the walking space. Intermediate destinations are the targets on pedestrian's way that can be accessible from the position that the pedestrian currently occupies. Each pedestrian is loaded to the system at her correspondent initial time. A reasonable speed is assigned to each pedestrian once she enters the system (initial speed). Initial direction is also assigned to the pedestrian in a way that she is almost heading towards her first intermediate destination. At each time step the utility of each alternative within pedestrian's choice set is calculated and afterwards the probability of choosing each alternative is obtained based on CNL model using Equation 3.27. The next step position of each pedestrian is selected to be the centre of the alternative with maximum probability. The position of each pedestrian is

updated at next time step. The speed and the direction of each pedestrian are also updated based on the successive positions of the pedestrian in the current and the next time step.

5.2 Data

In order to calibrate and validate the performance of the discrete choice model for pedestrians passing through bottlenecks and in high density situations some relevant data is required. The data used for this purpose in this study was extracted from video recordings of human movements at bottlenecks in two different locations.

5.2.1 Data Collection

The process of collecting and preparing data for each of these two locations is described in the following.

Duisburg Data, Germany

This data was collected at the campus Duisburg in Duisburg- Essen University, Germany. Kretz *et al.* (2006) collected the data to study pedestrian flow through bottlenecks. They conducted a controlled experiment with 94 participants who walked through a bottleneck with a depth of 40 centimetres trying to represent a normal walking situation (non-panic). The experiment repeated with ten different widths of the bottleneck including 40, 50, 60, 70, 80, 90, 100, 120, 140 and 160 centimetres. For each bottleneck width, multiple runs were performed (three runs for most of the bottleneck widths). The video was recorded from the top. A snapshot from the video recordings of a bottleneck with the width of 1 metre is shown in Figure 5.1.

This data is used in this study to measure the interpersonal distances between pedestrians in high density situations. As mentioned earlier in sub-section 2.3.2, an adequate space is required, both laterally and longitudinally, to avoid physical contacts. The size of personal space varies with density of the crowd; *i.e.,* as the density grows, the interpersonal distance decreases.

Figure 5. 1: A snapshot from German data

In order to measure interpersonal distances between pedestrians, snapshots from the recorded videos were created with the rate of two frames per second. The head coordinates of each pedestrian in every single image were stored, using Matlab and then were converted from pixel coordinates to metre coordinates. The distance between each pedestrian and the individuals in her surrounding environment in each frame was calculated using these coordinates. Interpersonal distance for each pedestrian is the minimum of all measured distances between her position and that of other pedestrians.

Moreover, results obtained from Kretz *et al.*'s experiments regarding the egress time and the flow of pedestrians passing through bottlenecks of different widths are compared with simulation results in this study for the sake of calibration and validation (sections 5.7and 5.8). In Kretz *et*

al.'s work, the egress time is defined as the time taken for the pedestrians, from the first to the last one, to pass a reference line (the white line on the ground, illustrated in Figure 5. 1). Flow is computed by dividing the total number of participants by the egress time. The specific flow is then computed for a unit of width (1 metre).

O-train Data, Ottawa

This data was collected at O-Train station at Carleton University, Ottawa, Canada during the times when the train stops at the station to load/unload passengers. Videos were recorded during November 2012 and purposefully in the morning rush hours (between 8 to 9 a.m) when the train is almost full of students heading to school. Therefore, when the train arrives at Carleton University the prevalent direction of movement is out-flow from the train. Almost no in-flows were observed. The camera was set on the fifth floor of one of the campus buildings (Minto Centre) which overlooks the station and provides a good view of southbound trains arriving at Carleton. Six videos were recorded. Video recordings started and ended with arrival and departure of the train and each of them was between 90 to 150 seconds long. A snapshot of the video is shown in Figure 5.2.

Figure 5. 2: A snapshot of O-Train data, Ottawa

The observed train has 3 exit doors with the width of 1.3 metres each. These exits are annotated in Figure 5.2 with letters A, B, and C. Each video was played in slow motion mode and the number of passengers passing through each door in every 5 seconds was counted. The egress time for each door was also measured. Egress time is defined as the total time from the first to the last passenger passing the door.

5.2.2 Results

The results obtained for both cases are reported in the following:

Duisburg Results

The interpersonal distances measured in different videos are reported in Table 5.1. The first column shows the name of the videos. The next four columns present the minimum and the percentile values of interpersonal distances.

Table 5.1: Interpersonal distances measured in different videos

Video name	Min (cm)	10th percentile	15th percentile	20th percentile
60_2	25.03	37.88	39.43	41.53
60_3	16.83	32.28	35.13	36.70
70_2	23.26	33.14	34.82	36.57
70_3	21.81	33.24	35.74	37.36
80_2	19.85	31.38	32.81	35.65
80_3	22.28	30.03	33.04	34.84
90_2	25.82	33.35	34.59	36.16
90_3	19.82	33.61	35.53	37.50
100-2	21.37	33.49	35.34	37.61
100-3	22.14	30.58	32.21	34.69
120_2	21.13	33.24	35.24	36.63
120_3	23.03	32.83	35.87	37.67
140_2	26.40	33.82	36.06	37.51
140_3	24.50	32.86	35.67	36.89
160_2	23.92	33.56	36.75	38.35
160_3	21.97	33.25	36.38	38.31
average	22.45	33.03	35.29	37.12

Table 5.2 provides the observed interpersonal distances with regard to the angle between the pedestrian's direction of movement and the position of nearby pedestrians. Angle 90° shows the lateral distance and angle 0° shows the longitudinal distance that the pedestrian maintains with other pedestrians. Each value reported in this table is the average of the values obtained from different videos. The first column shows different angle ranges from 0° to 90°. The second column is the number of observations for each category of angle range. Minimum and percentile values are shown in subsequent columns.

Table 5.2: Interpersonal distances with respect to the angle between the direction of movement and position of other pedestrian

angles (degree)	Num of Observations	Min (cm)	10th percentile	15th percentile	20th percentile
0-5	281	27.69	30.77	33.17	35.23
5-10	259	31.78	34.81	36.49	38.20
10-15	283	30.69	34.88	36.79	38.80
15-20	254	30.82	35.68	37.34	38.86
20-25	309	30.81	36.08	38.75	41.13
25-30	343	32.58	36.42	38.44	40.22
30-35	336	31.84	36.12	38.45	40.33
35-40	382	31.09	36.78	39.06	41.43
40-45	332	32.73	37.70	40.15	41.99
45-50	359	31.11	36.78	38.94	40.85
50-55	302	31.23	35.07	36.88	39.26
55-60	245	31.37	33.99	36.33	38.69
60-65	214	32.36	34.64	36.07	37.25
65-70	191	34.05	35.91	37.49	39.32
70-75	214	33.70	36.27	37.39	38.04
75-80	216	29.15	32.41	34.10	35.38
80-85	195	29.38	32.57	35.42	36.76
85-90	215	31.05	33.14	34.91	36.18

Looking at the values in the tables above, one can conclude that in high density situations interpersonal distance is not a function of the angle between pedestrians, that is, the distance a pedestrian keeps with other nearby people does not depend on whether they are arranged in lateral or longitudinal directions. A value between 35 to 40 centimetres can be selected as a distance threshold between pedestrians based on 15th and 20th percentile values in Table 5.1 and Table 5.2. To be on the safe side, here in this study, the value of 40 centimetres has been considered as the distance threshold in high density situations.

The egress time and the flow computed by Kretz *et al.* in different experiments are also reported in Table 5.3 and Table 5.4 . Some of the experiments were conducted with 80 participants and some others with 100 participants.

Table 5.3: Measured egress time and flow for different bottlenecks (80 participants)

80 Participants			
Bottleneck Width (cm)	**Egress Time Range (sec)**	**Flow Range (ped/s)**	**Special Flow Range (ped/m.s)**
40	89-90	0.85-0.9	2.2-2.25
50	76.5-78	1-1.05	2-2.1
60	69-73.5	1-1.15	1.75-1.9
70	63.5-65.5	1.2-1.25	1.7-1.8
80	54-57.5	1.35-1.45	1.7-1.85
90	46-48	1.65-1.7	1.8-1.9
100	39-47	1.65-2.05	1.65-2.05
120	31-42	1.85-2.55	1.55-2.15
140	35.5-38	2.05-2.2	1.45-1.6
160	33-37.5	2.1-2.4	1.3-1.5

Table 5.4: Measured egress time and flow for different bottlenecks (100 participants)

100 Participants			
Bottleneck Width (cm)	**Egress Time Range (sec)**	**Flow Range (ped/s)**	**Special Flow Range (ped/m.s)**
40	111.5-113	0.85-0.9	2.15-2.25
50	96-98.5	1-1.05	2-2.1
60	88-88.5	1.1-1.15	1.85-1.9
70	79-81.5	1.2-1.25	1.7-1.8
80	67.5-71.5	1.35-1.5	1.7-1.85
90	60-65.5	1.5-1.65	1.65-1.85
100	49-58.5	1.7-2.1	1.7-2.1
120	43-49	2-2.3	1.65-1.95
140	44.5-50	1.95-2.3	1.4-1.6
160	46-47.5	2.05-2.15	1.3-1.35

O-train Results

Table 5.5 reports the egress time and the number of passengers passing through each door of the train.

Table 5.5: Egress time and number of passengers passing each door of the train

Door A						
Number of passengers	45	34	20	39	68	13
Egress time (sec)	48	28	16	31	51	9
Door B						
Number of passengers	53	54	40	68	80	29
Egress time (sec)	56	36	28	46	58	20
Door C						
Number of passengers	77	27	23	47	67	18
Egress time (sec)	55	17	18	32	48	14

Figure 5.3 shows the relation between the total number of passengers leaving the train and the egress time. Expectedly, egress time increases as the number of passengers increases. A line has been fitted to data and the results of ANOVA test implies that there is a significant relationship between two variables which can be well explained by this line (shown in Figure 5.3). The

obtained value of R-square is 0.83. Therefore, a linear relationship between these two variables has been found.

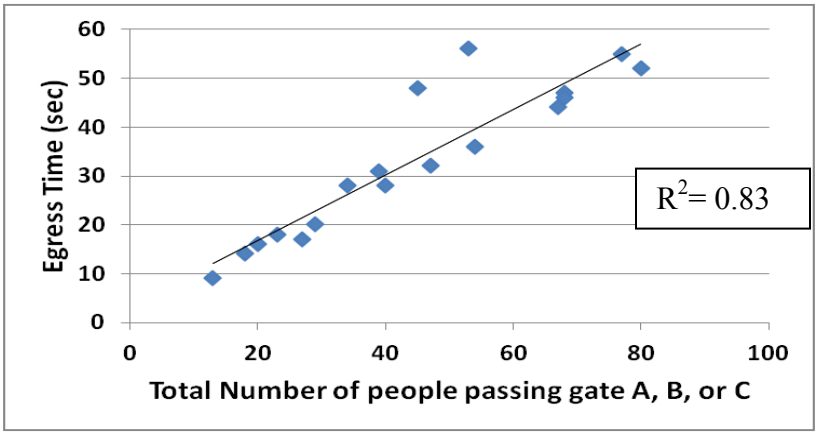


Figure 5.3: Relation between number of people passing the door and egress time

Figure 5.4 displays the observed flows passing through doors A, B, or C for the time interval of 5 seconds. The maximum and average flows are 7.7 *ped/m* and 5.17 *ped/m* respectively with the most frequent observed value of 6.15 *ped/m* for the time interval of 5 seconds. Therefore, the maximum and mode flows are 1.54 *ped/m.s* and 1.23 *ped/m.s* respectively.

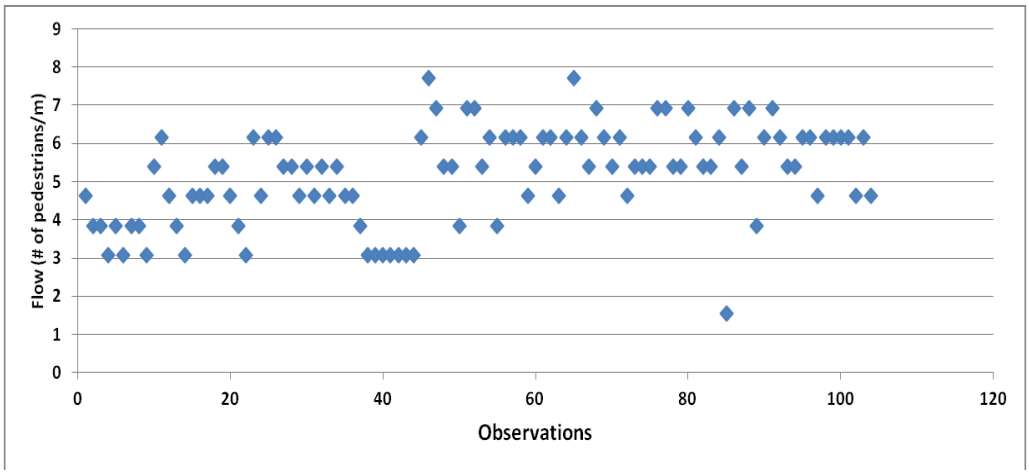


Figure 5.4: Observed flows every 5 seconds

5.3 Simulating Pedestrians Movement at Bottlenecks in High Density Situations Using Next Step Model

In order to simulate Kretz *et al.*'s experiments a rectangular walking area of 9 by 4 metres has been built in the simulation tool and the walls have been defined as obstacles. Therefore, when the line that connects the pedestrian's current position to the centre of the alternative crosses any wall, the alternative is considered to be unavailable in the simulation tool. One hundred pedestrians have been loaded to the system at time step zero. They are placed in 5 rows with each row containing 20 pedestrians. Their initial positions have been selected in a way that they maintain at least 40 centimetres distance from nearby pedestrians. In this study, the initial positions, speeds, directions, and destinations of pedestrians have been chosen to be deterministic and do not change for different simulation runs, so as to obtain comparable results from different model specifications.

A very slow initial speed (0.01 *m/s*) has been assigned to all pedestrians (except for 3 of them) as they are waiting to pass the door. Next step model cannot predict the next position of pedestrians with the speed of zero (decision-maker cannot be a stationary individual) and that is the reason why a small speed value has been assigned to pedestrians in the system. Only the 3 pedestrians in front of the door have higher initial speeds. The pedestrian right in front of the door has initial speed of 1.2 *m/s* and the two other pedestrians on the left and right of her have the speed of 0.6 *m/s*. All pedestrians have a final destination and an intermediate destination. Their final destinations are randomly set outside the defined rectangular walking area. The intermediate destinations of all pedestrians are located at the centre of the door at its outside edge. A simulated position of pedestrians at a random time step for a bottleneck of 1 metre wide is shown in Figure 5.5. Each cross sign represents a pedestrian and the numbers shown in the figure are

pedestrians' IDs. The sample trajectories of two pedestrians are also shown in Figure 5.6 and the trajectories of all pedestrians are displayed in Figure 5.7.

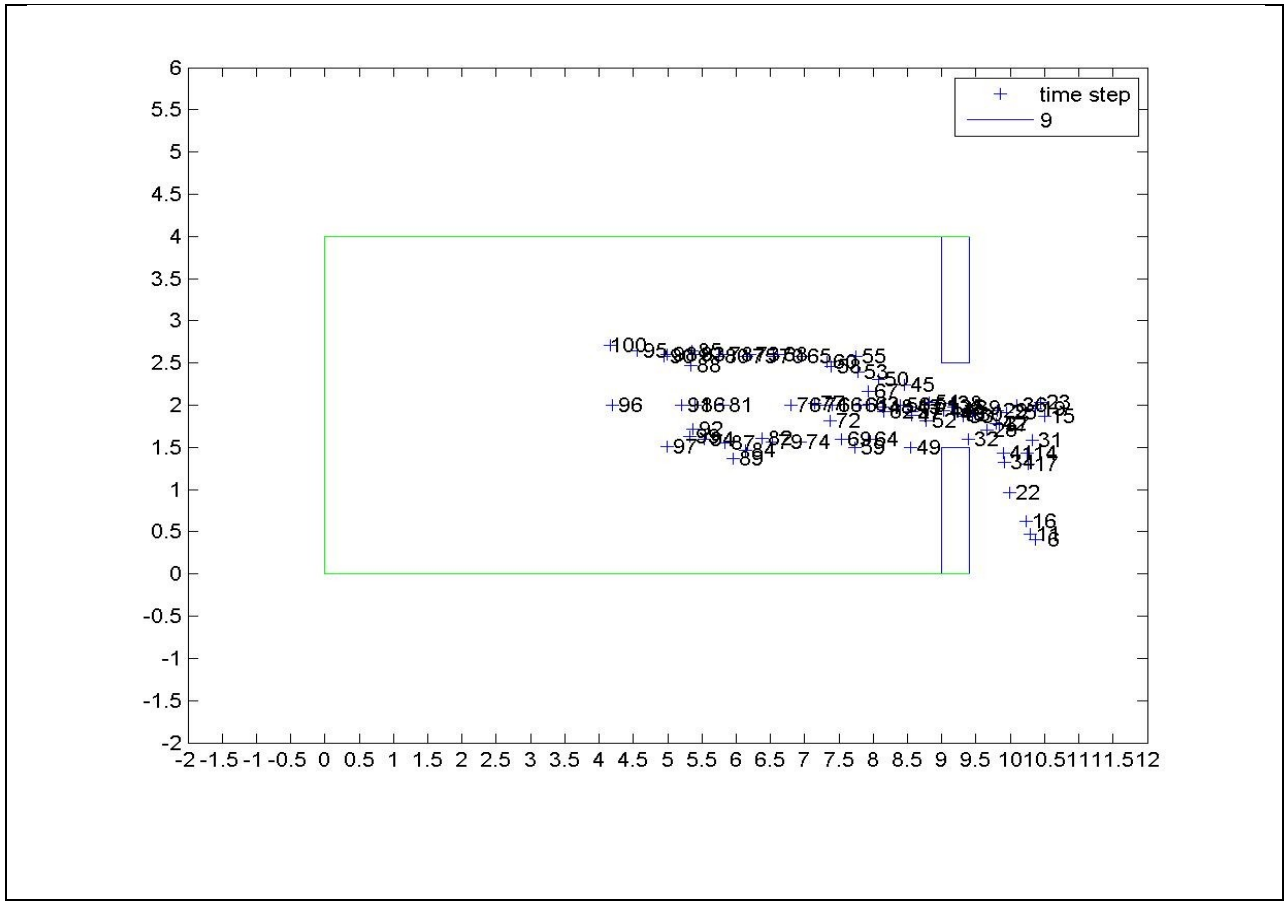


Figure 5.5: A simulated frame of pedestrian movements at bottleneck based on next step model (axes units are metre)

Interpersonal distances between pedestrians have been measured at every time step. The distances between the pedestrian and those individuals in front of her (the angle between the pedestrian's movement direction and the position of other individuals is between 0° to 90°) have been calculated and finally the interpersonal distance of each pedestrian is defined as the minimum of all these distances. Figure 5.8 shows the frequency of interpersonal distances. Seventy eight percent of interpersonal distances are less than the distance threshold (0.4 m).

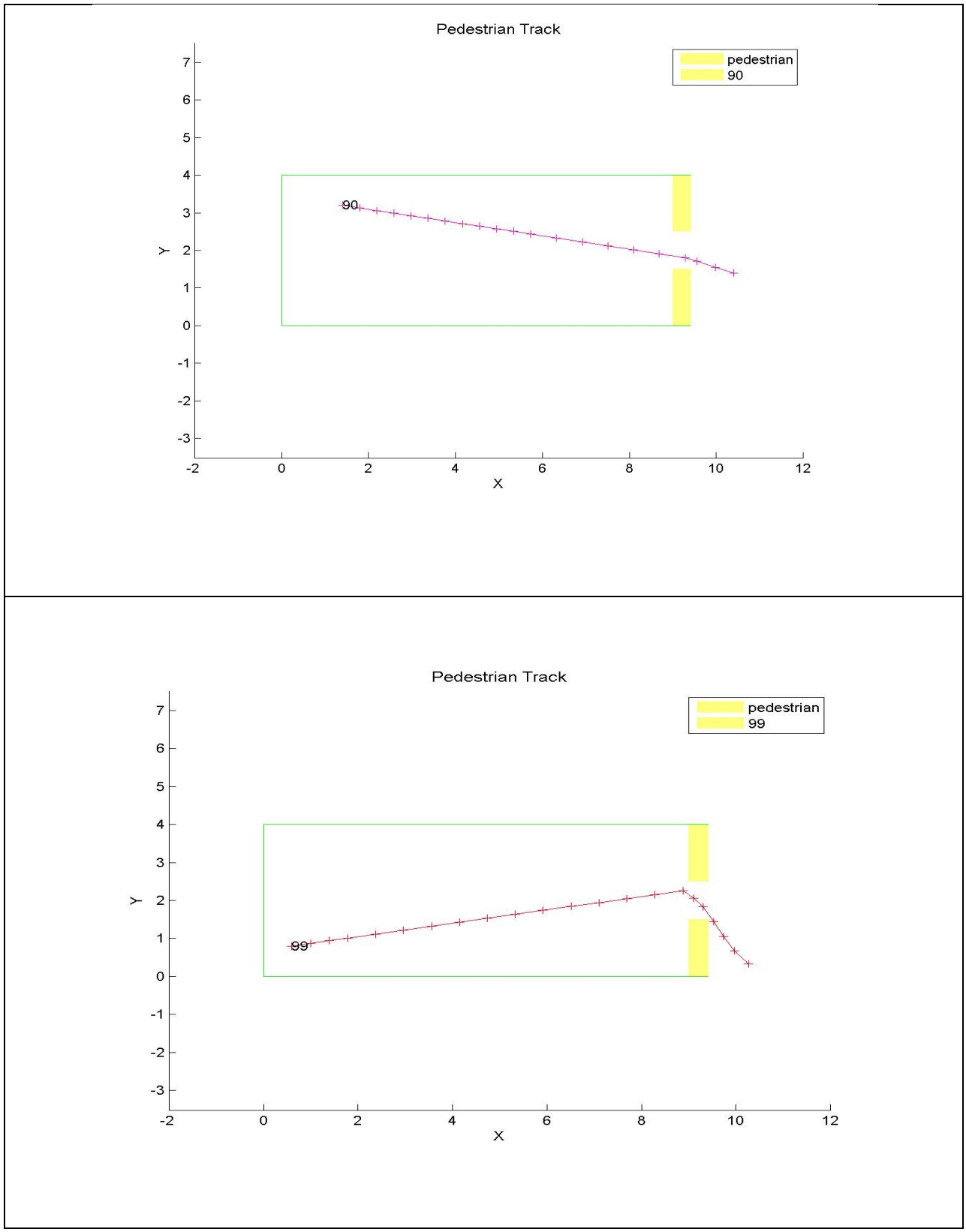


Figure 5.6: Samples of simulated pedestrian trajectories based on next step model (axes units are metre)

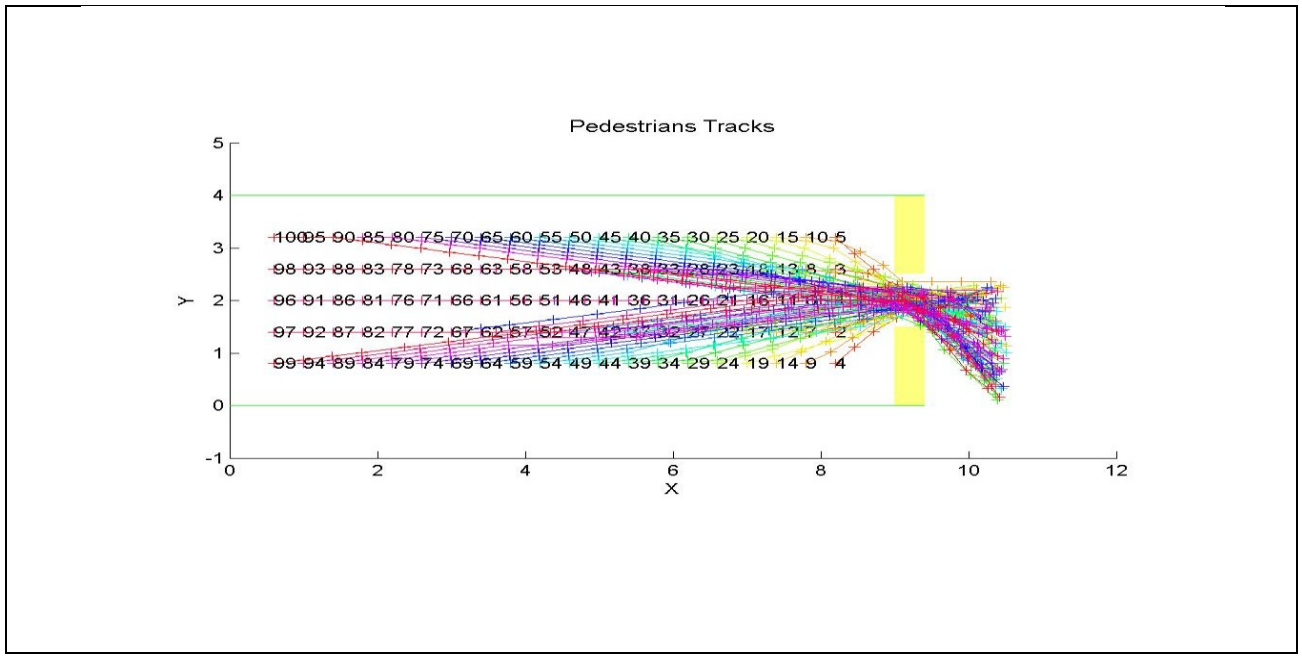


Figure 5.7: Trajectories of all pedestrians based on next step model (axes units are metre)

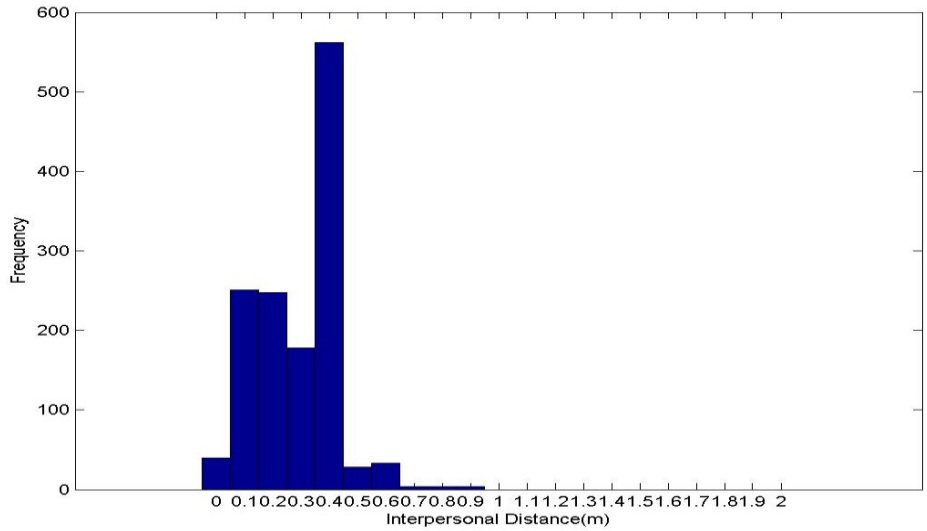


Figure 5.8: Histogram of interpersonal distances (next step model)

Smooth motion trajectories such as those illustrated in Figure 5.6 and Figure 5.7 have been obtained from next step model. The time it takes for all one hundred pedestrians to pass through the bottleneck (egress time) is 8.5 seconds which is not reasonable; it is much less than the egress

time reported by Kretz *et al.* (Table 5.4). Also, it was found that on average 13 pedestrians pass the door of 1 metre wide in one second (average flow), which is not realistic. The main reason for this occurrence is that pedestrians conflict with each other while passing the door and do not maintain adequate interpersonal distance with nearby individuals as shown in Figure 5.5. Based on the collision avoidance behaviour in next step model, each pedestrian avoids collision with potential colliders in the environment while walking. However, in this model colliders are only limited to those pedestrians that walk almost in the opposite direction of the decision-maker. In some conditions such as high density situations this brings about problems. Sometimes, pedestrians walking in nearly the same direction can also collide or get so close to each other that the safe interpersonal distance between them could not be maintained. Also, colliders are only limited to moving pedestrians; Fixed obstacles such as walls or columns have not been considered in the model. Next step model needs to be developed by incorporating the aforementioned effects.

The results obtained above indicate that next step model does not perform well when pedestrians are supposed to walk through bottlenecks at high density situations and has some limitations. Therefore, some modifications have been suggested in this research to improve the model's functionality in discussed situations. In the remaining sections of this chapter, it is tried to improve next step model by overcoming some of these limitations.

5.4 Model Development

Two behaviours have been introduced to develop next step model in order to capture the behaviour of pedestrians at bottlenecks and in high density situations. These two behaviours are:

maintain interpersonal distance behaviour and wall avoidance behaviour. In the following these behaviours are explained.

5.4.1 Maintain Interpersonal Distance Behaviour

This behaviour is introduced to allow pedestrians to maintain adequate distance from each other. The unidirectional pedestrian movement at bottleneck and in high density situation is of interest in this study. Most of the potential conflicts in mentioned situation occur between pedestrians walking close to each other moving in the same direction. Therefore, it is assumed that decision-makers avoid pedestrians with the similar moving directions as well. The model was run with various specifications and its performance was closely examined to find the best specification. Finally after many trials, maintain interpersonal distance behaviour was modeled based on buffering and blockage concepts. It is assumed that decision-makers avoid pedestrians in the nearby environment even if they are relatively moving in the same directions (buffering). Moreover, it is assumed that pedestrians do not choose alternatives with a distance less than the threshold to the projected position of nearby individuals (blockage). The distance threshold in high density situations is considered to be 40 centimetres.

Similar to collision avoidance behaviour in next step model, maintain interpersonal distance behaviour is studied in a region of interest. The region of interest is a semi-circle in front of each pedestrian with the radius that is 5 times bigger than the radius of the choice set. The focus of this behaviour is on maintaining a proper distance with nearby pedestrians. Therefore, a smaller region of interest has been selected compared to that of collision avoidance behaviour. Clearly, pedestrians moving in opposite direction affect decision-maker's behaviour in a longer range.

Figure 5.9 illustrates the elements of this behaviour. The region of interest is displayed with dashed lines. The pedestrians in this region are identified as potential colliders (k) and are shown with grey circles. Blue circles are the projected position of these pedestrians in next time step. Similar to group behaviour in chapter 4, it is assumed that the decision-maker predicts each potential collider's future position based on the assumption of constant velocity. Hence the projected position of each pedestrian in next time step is expected to be cell 17 (see Figure 3.4). D_k is the distance between centre of the alternative and the projected position of the potential collider. If there is more than one potential collider in the region of interest, the collider for each alternative is selected among all potential colliders whose projected position has the shortest distance to the centre of the alternative.

The decision-maker's behaviour to maintain interpersonal distance when the distance of the collider to the centre of the alternative is larger than the distance threshold is captured by the following term (buffering):

$$I_{ip}\beta_{ip}e^{(D_{ip}-D_{th})\rho_{ip}} \qquad \text{(Equation 5. 1)}$$

where:

- I_{ip} is equal to one if a collider is identified in the region of interest and zero otherwise,
- D_{ip} is the distance of the collider to the centre of the alternative. This is further illustrated in Figure 5. 9,
- D_{th} is the distance threshold and as mentioned before it is equal to 40 centimetres, and
- β_{ip} and ρ_{ip} have to be estimated; negative signs are expected for both of them.

In order to be in accordance with collision avoidance behaviour, the effect of distance on the utility of this behaviour is modeled, using exponential decay function. The attractiveness of the alternative decreases as the distance of the collider to the centre of the alternative decreases.

The alternative is considered to be unavailable in case the distance of the collider to the centre of the alternative is lower than the distance threshold.

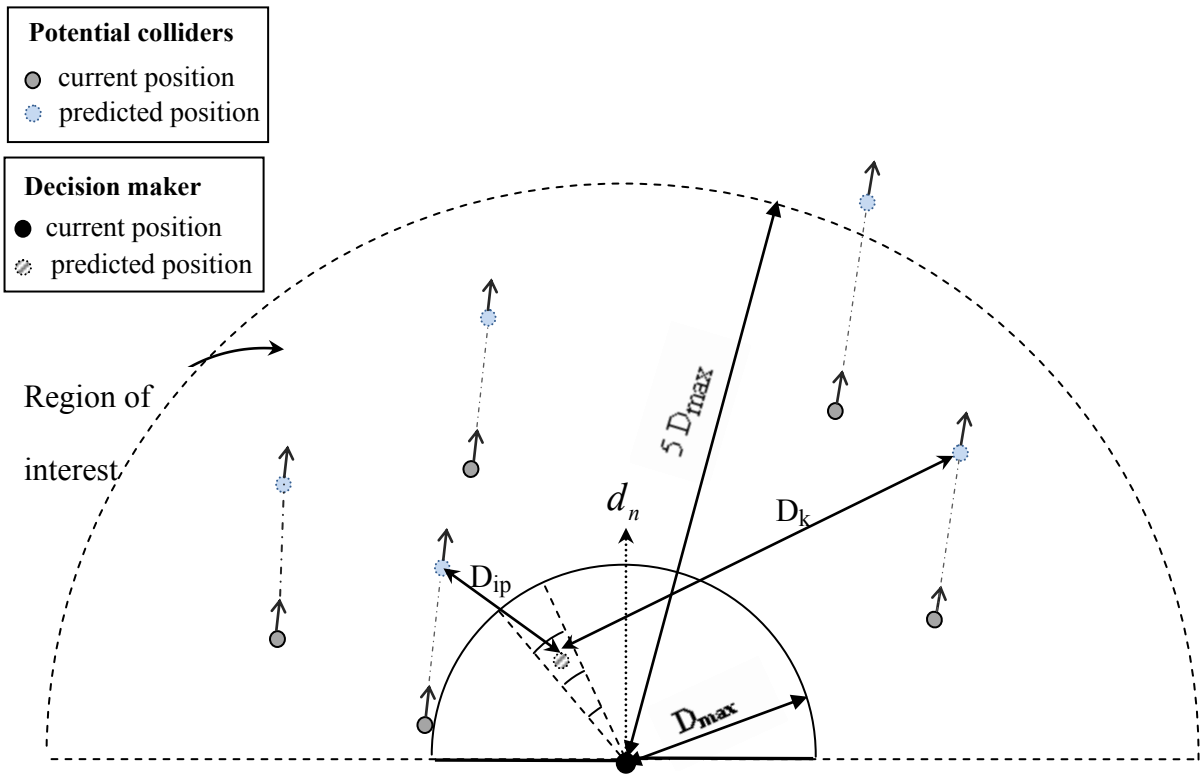


Figure 5. 9: Elements of maintain interpersonal distance behaviour

5.4.2 Wall Avoidance Behaviour

This behaviour captures pedestrian's tendency to maintain a distance from walls and not to get too close to the walls. The model specification for this behaviour was inspired by collision avoidance behaviour in next step model. If there is any wall in a specific region of interest, it is assumed that the decision-maker tries to avoid it. The attractiveness of alternatives varies based

on their distance to the wall. The region of interest for each of the 11 radial cones (Figure 3.2) is restricted to the left and right direction boundaries of the radial cone (d_l and d_r). The distance between the centre of the alternative and the wall has to be less or equal to 5 times the radius of the choice set (D_{max}). Figure 5.10 shows the elements of this behaviour. The smaller region of interest compared to that of collision avoidance behaviour is due to the fact that pedestrians get closer to walls rather than to those pedestrians walking in opposite directions.

The utility terms associated with this behaviour are as follows:

$$I_w \beta_w e^{D_w \rho_w} \qquad \text{(Equation 5. 2)}$$

where:

- I_w is equal to one if any wall is present in the alternative's region of interest and zero otherwise,
- D_w is the shortest distance between the centre of the alternative and the wall (or walls). This is further illustrated in Figure 5.10, and
- β_w and ρ_w have to be estimated; negative signs are expected for both of them.

The attractiveness of the alternative decreases as its distance to the wall decreases. If the line that connects the current position of the pedestrian to the alternative (potential position of pedestrian in next time step) intersects with any present wall, the alternative is considered to be unavailable.

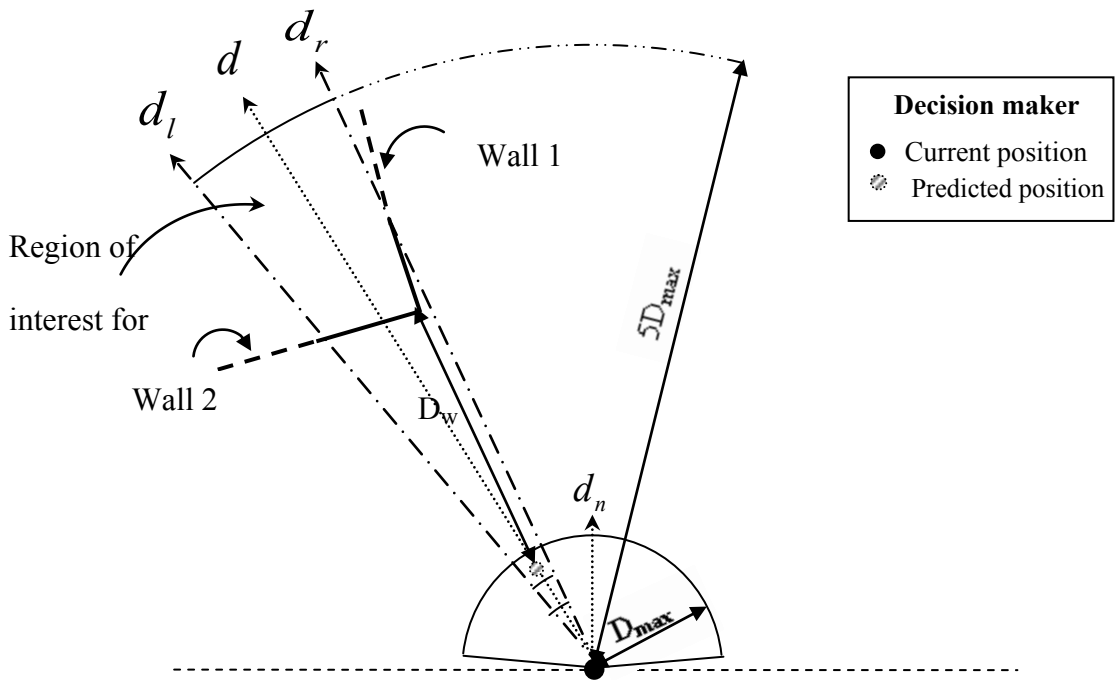


Figure 5.10: Elements of wall avoidance behaviour

5.5 Model Revisions

Running many trials of the simulation model with the two newly presented utility terms showed that some changes should be made to model specification in order to improve its functionality in high density bottlenecks.

It was observed that pedestrians sometimes move around rather than walking straight to their destinations (see Figure 5.11). Wandering around the walking area (changing movement direction) is expectable to some extent in high density situations. However, when this phenomenon occurs frequently and more than what is naturally expected it becomes a matter of concern.

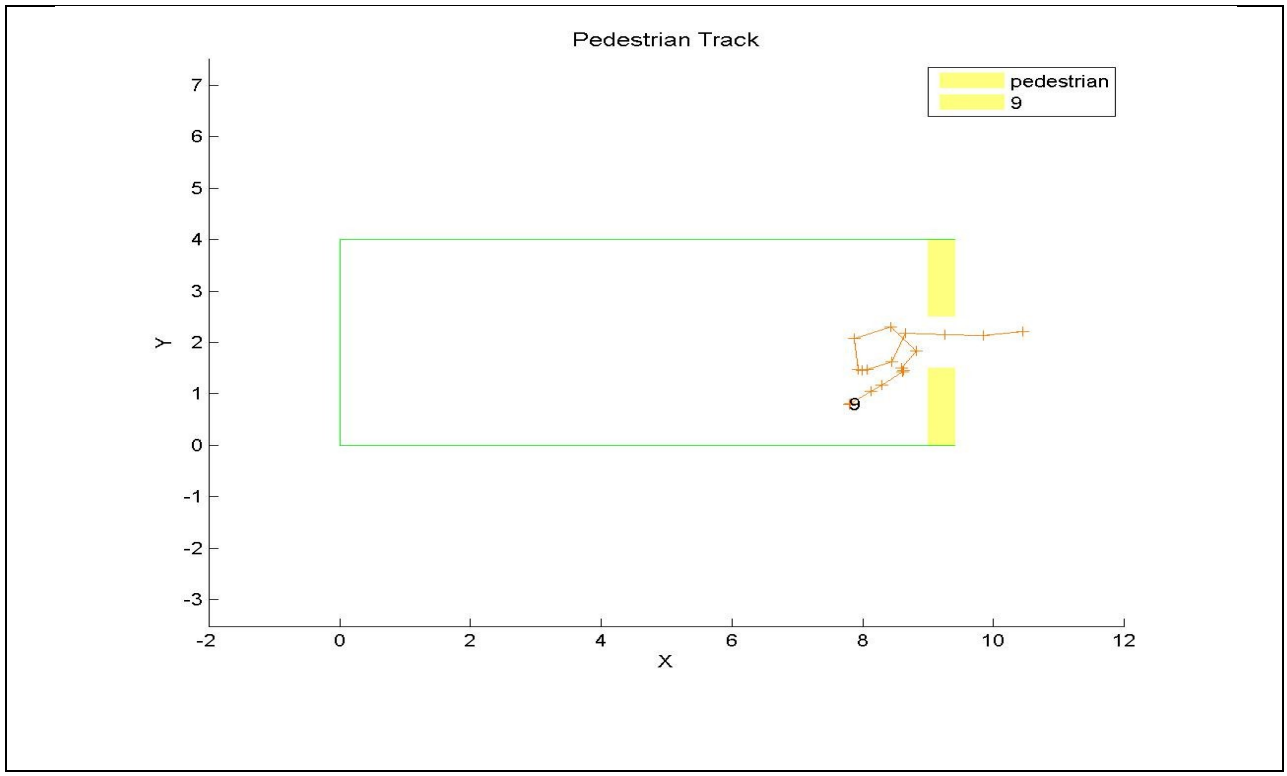


Figure 5.11: Example of wandering around in trajectory of pedestrians (axes units are metre)

To fix this problem, a new row of deceleration alternatives have been added to the choice set to provide each decision-maker with the possibility of choosing an alternative with a very low speed (close to zero). Since pedestrians' behaviours in high density situation is of interest, it is assumed that when the area in front of a decision-maker is crowded and is not free of pedestrians or obstacles, the decision-maker sometimes prefers to walk in a very low speed (or not to move at all) and wait till upcoming time steps when she has more freedom of motion towards her destination, rather than choosing an alternative which is clearly out of her destination route; the one which makes her path longer and can lead to the occurrence of wandering movements. Therefore, the choice set consists of 44 cells in this study with 22 alternatives belonging to deceleration zone (and deceleration nest) and 11 alternatives in each of constant speed and acceleration zones. For a time step t, the newly added deceleration row ranges from 0 to $0.25v_n t$

with its centre being at $0.125v_n t$. The structure of the choice set is illustrated in Figure 5.12. Having 44 cells instead of 33 cells mitigates the level of wandering in pedestrians' trajectories.

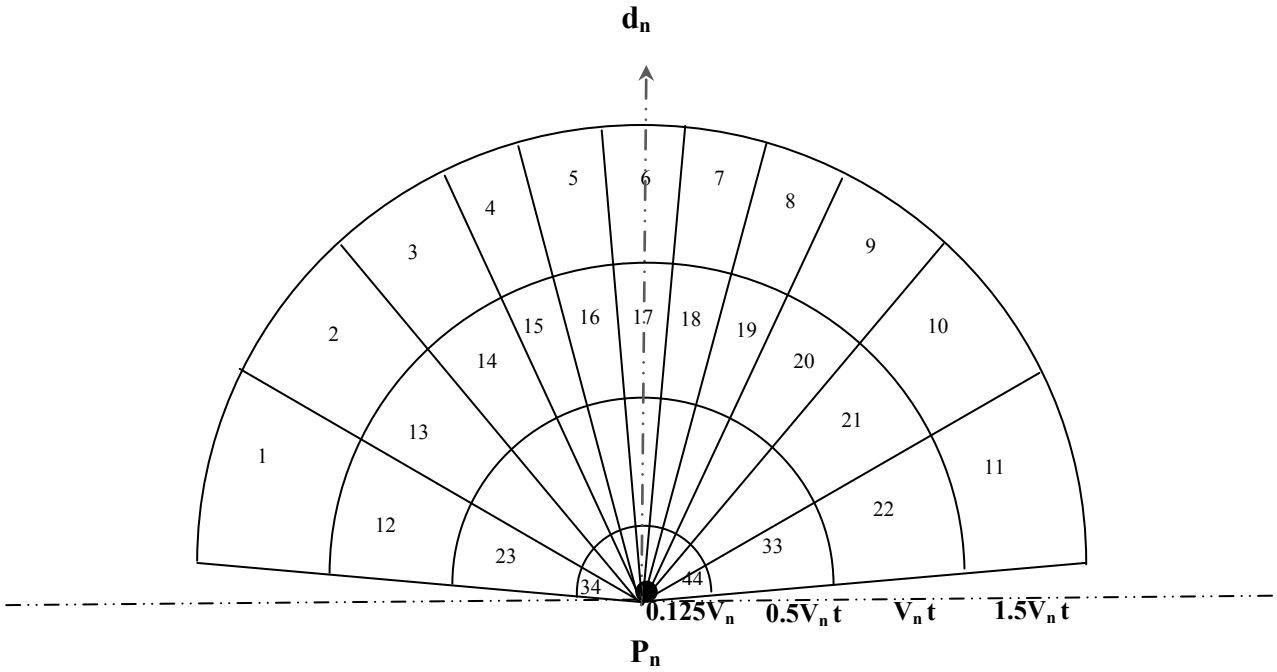


Figure 5.12: Revised choice set consisting of 44 cells

Second, it turned out that the sensitivity of pedestrian's decision making to the utility terms that captures *toward destination* behaviour should increase. This behaviour deals with the decision-maker's tendency to decrease the distance and direction to the destination. Next step model was calibrated based on pedestrian movements on crosswalks where the pedestrians' destinations are usually any point on the other side of the crosswalk, reaching which does not require severe changes in movement directions. Therefore, in most cases toward destination behaviour of pedestrians was better explained through *keep direction* behaviour and consequently estimation results obtained for toward destination may not be compatible with other situations and cannot be used interchangeably in other circumstances. Many of the pedestrians who are moving through

bottlenecks need to change their movement directions to reach their intermediate destinations which are located at the centre of the door (bottleneck). Running multiple simulation trials showed that the sensitivity to toward destination behaviour should be improved. This reduces the level of wandering in individuals' trajectories as well.

5.6 Keep Simulation on Track

In addition to the changes made to the model specification, some conditions are required to keep the simulation on track. All these considered criteria are discussed in the following:

Criterion 1: if all the alternatives a pedestrian can choose in time step t are blocked with walls (obstacles), the pedestrian does not select any alternative in that time step. Instead, she turns around (her movement direction changes 180 degrees). Therefore, in time step $t+1$ her position and speed are the same as those of last time step, but her movement direction changes 180° and then she continues to move from this point.

Criterion 2: if all the alternatives that a pedestrian can choose in time step t (except for those that are blocked with walls) are blocked with the presence of other pedestrians, *i.e.,* alternatives are unavailable, the pedestrian does not select any alternative in that time step. Instead, she stays in her position and then in time step $t+1$ she continues to walk (if there is any available alternative to choose) with the same direction and speed as those of last time step.

Criterion 3: In some circumstances the simulation freezes. This happens because a group of pedestrians block each other in a way that none of them can move, *i.e,* the distance of all alternatives to the projected position of pedestrians around is less than the distance threshold. This problem would not be resolved unless some criteria are taken into account in the simulation

process. Therefore, when some of the pedestrians block each others' ways, the right of way is given to that blocked pedestrian who is closer to her destination (intermediate or final destination) and other blocked pedestrians do not move. By doing this, all blocked pedestrians gradually move in upcoming time steps. This keeps the simulation on track and resolves the freezing problem.

Figure 5.13 shows a simulated frame in which pedestrians block each others' way near the exit. The arching pattern around the bottleneck (door) which occurs in high density situations is clearly observable in this figure. This problem was fixed with the giving way criterion as discussed above.

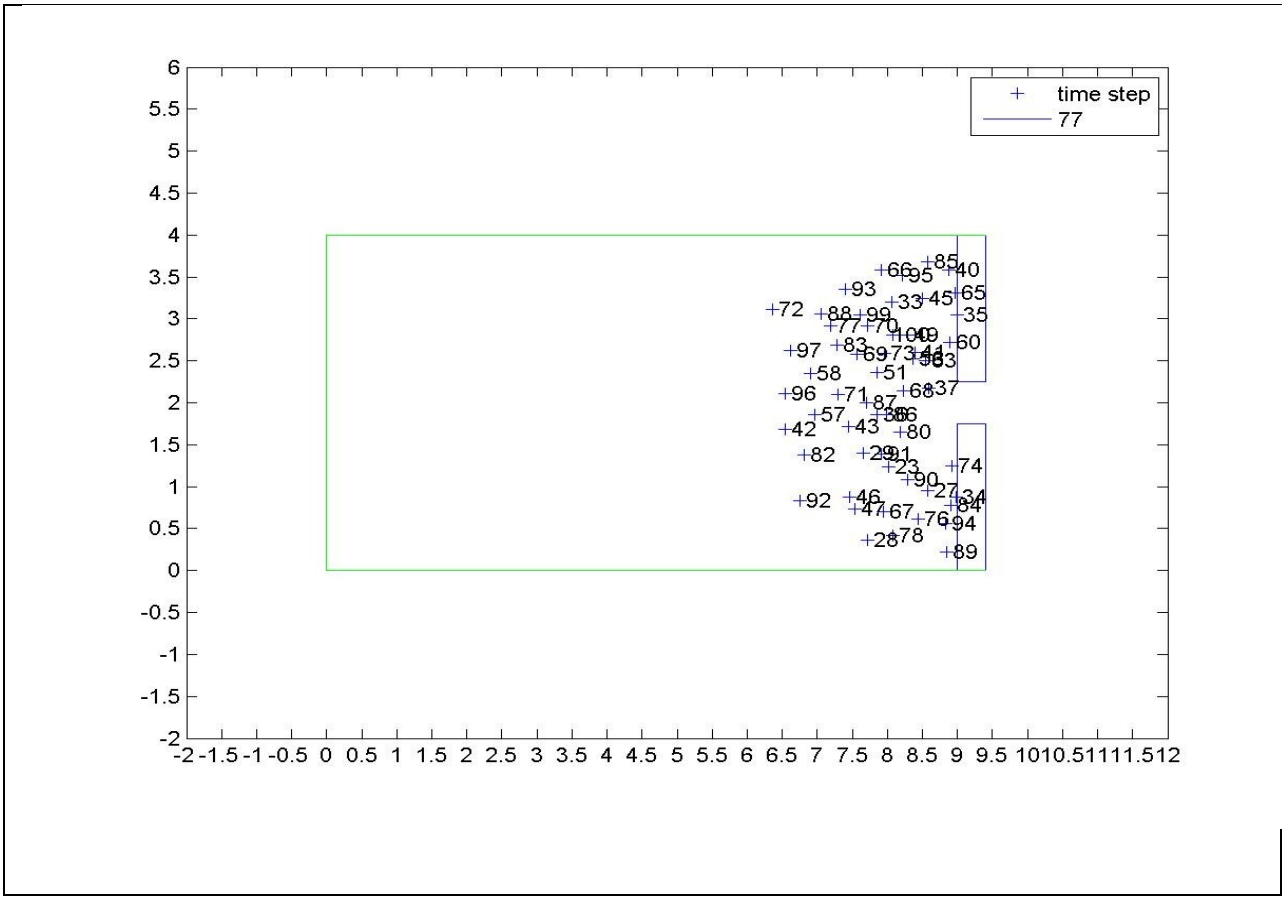


Figure 5.13: Observed arching pattern near the exit (axes units are metre)

5.7 Model Calibration

Calibration and validation of microscopic pedestrian models can be performed using quantitative or qualitative analysis. In quantitative analysis, measurable model parameters such as flow or egress time are compared with real data. However, in qualitative analysis, patterns that are formed in simulation model are compared with those emerging patterns that are actually observed in real pedestrian crowds such as lane formations or oscillations at bottlenecks. Simulated visuals have to be checked to determine whether the animations are realistic and acceptable or not. In this study, the model has been calibrated and validated by reproducing macroscopic quantities of pedestrian dynamics. Moreover, Trajectories of pedestrians are checked to see whether they represent coherent and realistic pedestrian movements or not.

There are 6 parameters that have to be estimated based on the developed model specification discussed above. These parameters are β_{ddist}, β_{ddir}, β_{ip}, ρ_{ip}, β_w, and ρ_w. These parameters are defined in section 5.4 and sub-section 3.2.2. After running the simulation model with different potential values for these 6 parameters, 5 values for each of β_{ip}, ρ_{ip}, β_w, and ρ_w and 10 values for parameters β_{ddist} and β_{ddir} have been selected for the sake of calibration. The number of possible combinations among the parameters that have to be estimated plays an important role in calibration process. If this number is very large, the calibration process cannot be performed in a reasonable time. In the current case, there are more than 15625 (considering 5 values for each parameter which results in 5^6) possible combinations. This value is so large. Therefore, due to the limitation of the number of runs, it is not possible to estimate all these parameters at once. 4 parameters of β_{ip}, ρ_{ip}, β_w, and ρ_w were fixed and the model was run with different values of

β_{ddist} and β_{ddir}. Then, the best obtained result has been chosen in order to estimate the other 4 parameters.

The model has been calibrated based on one of the experiments conducted in Kretz *et al.*'s (2006) work in which 100 pedestrians pass through a bottleneck of 1 metre wide.

5.7.1 Calibration of Toward Destination Behaviour

10 different values have been considered for β_{ddist} and β_{ddir} for the sake of calibration:

β_{ddist}= [-1.5, -2, -3, -4, -5, -6, -7, -8, -9, -10]

β_{ddir}= [-0.0793, -0.1, -0.15, -0.2, -0.3, -0.4, -0.5, -0.6, -0.7, -0.8]

β_{ddir}=-0.0793 and β_{ddist}=-1.5 are values obtained from next step model and have been used in the estimation process for the sake of comparison. The values of β_{ip}, ρ_{ip}, β_w, and $\rho_w$have been fixed to -5, -5, -5, and -8 respectively. The total of 100 (10^2) simulation runs have been performed. Among all 100 combinations of parameters, an optimal parameter set has to be selected in a way that provides the best match to real pedestrian data.

The measures of performance selected for the calibration process are egress time (flow), total walking distance, and the number of incoherent trajectories. Egress time is the time from the first to the last pedestrian passing the bottleneck. The flow through bottleneck is calculated by dividing the total number of pedestrians to the egress time in order to be consistent with Kretz *et al.* 's (2006) calculation. The egress time obtained from the simulation is compared to the values reported by Kretz *et al.* (2006). According to Kretz *et al.*'s results for the bottleneck of 1 metre wide, acceptable ranges for egress time are determined to be 45 to 60 seconds (or flow values

between 1.6 to 2.2 *ped/s*). The egress time and flow obtained for 40 combinations of parameters (out of 100 combinations) are presented in Table 5.6. Those parameter sets with acceptable egress time (or flow) are selected as candidate parameter sets.

The total walking distance and the number of incoherent trajectories were measured for all candidate sets. The results for some of the candidate sets are shown in Table 5.7. Total walking distance is the summation of all pedestrians' trajectory lengths. In high density situations when all pedestrians aim to pass a single point, presence of some non-smooth and incoherent trajectories are natural. Incoherent trajectories are referred to those trajectories in which there are signs of wandering around. Having a utility function that is more sensitive to walking toward destination behaviour decreases the level of wandering in trajectories by giving the decision-maker a stronger will to walk straight towards her destination. The better parameter set is the one that results in a fewer number of incoherent trajectories. All pedestrians' trajectories have been checked one by one in order to measure this quantity (visualization check). The smaller value of total walking distance presents the better parameter set as well. When this value is bigger, it means it takes more time for pedestrians to reach their destinations, which itself is an evidence of occurrence of more wandering movements in the walking area.

Table 5.6: Sample of simulated egress time for different parameter sets regarding toward destination behaviour

β_{ddir}	β_{ddist}	Egress time (*sec*)	Flow (*ped/s*)
-0.0793	-1.5	58	1.72
-0.0793	-3	47	2.13
-0.0793	-4	69	1.45
-0.0793	-9	37.5	2.67
-0.1	-2	64.5	1.55
-0.1	-3	55	1.82
-0.1	-5	47	2.13
-0.1	-7	44.5	2.25
-0.15	-5	66.5	1.50
-0.15	-6	50.5	1.98
-0.15	-8	45.5	2.20
-0.15	-10	51	1.96
-0.2	-1.5	71.5	1.40
-0.2	-2	64	1.56
-0.2	-5	51	1.96
-0.2	-9	46.5	2.15
-0.3	-1.5	87	1.15
-0.3	-2	67.5	1.48
-0.3	-3	59.5	1.68
-0.3	-5	56.5	1.77
-0.4	-1.5	74	1.35
-0.4	-2	76	1.32
-0.4	-4	55.5	1.80
-0.4	-9	52.5	1.90
-0.5	-1.5	84.5	1.18
-0.5	-3	64	1.56
-0.5	-8	48.5	2.06
-0.5	-9	51.5	1.94
-0.6	-3	65	1.54
-0.6	-4	64.5	1.55
-0.6	-5	53.5	1.87
-0.6	-10	59	1.69
-0.7	-1.5	79.5	1.26
-0.7	-5	77.5	1.29
-0.7	-8	58.5	1.71
-0.7	-10	49.5	2.02
-0.8	-2	77.5	1.29
-0.8	-5	58.5	1.71
-0.8	-7	59	1.69
-0.8	-9	48.5	2.06

Table 5.7: Evaluation of candidate parameter sets regarding toward destination behaviour

β_{ddir}	β_{ddist}	Egress time (*sec*)	Flow (*ped/s*)	Total walking distance (*m*)	Number of incoherent trajectories (out of 100)
-0.0793	-1.5	58	1.72	941.34	23
-0.0793	-2	56.5	1.77	1026.50	33
-0.0793	-5	52	1.92	973.32	33
-0.0793	-10	48.5	2.06	922.60	19
-0.1	-3	55	1.82	902.21	25
-0.1	-6	49.5	2.02	887.81	21
-0.1	-8	48	2.08	860.63	24
-0.1	-9	54.5	1.83	828.62	16
-0.15	-6	50.5	1.98	905.33	21
-0.15	-9	57.5	1.74	822.78	20
-0.15	-10	51	1.96	849.00	20
-0.2	-3	57	1.75	803.00	14
-0.2	-4	60	1.67	868.31	21
-0.2	-5	51	1.96	825.24	19
-0.2	-6	50	2.00	829.46	16
-0.2	-8	50	2.00	833.22	17
-0.2	-10	50.5	1.98	833.35	18
-0.3	-5	56.5	1.77	836.85	14
-0.4	-4	55.5	1.80	882.19	18
-0.5	-8	48.5	2.06	866.83	22
-0.5	-9	51.5	1.94	839.85	20
-0.6	-10	59	1.69	860.28	20
-0.7	-10	49.5	2.02	841.10	14
-0.8	-7	59	1.69	821.65	15
-0.8	-9	48.5	2.06	857.43	17
-0.8	-10	57	1.75	814.75	15

The row shown in yellow in Table 5.7 is the parameter set obtained from next step model (the default set). The candidate parameter sets have been evaluated based on the total walking distance and the total number of incoherent trajectories and the best parameter set was selected. Compared to the default set, smaller values for the total walking distance and the number of

incoherent trajectories have been obtained for most of other candidate sets, as shown in Table 5.7. The parameter set with β_{ddist} equal to -3 and β_{ddir} equal to -0.2 has been chosen as the best parameter set based on its walking distance and animations (shown in purple in Table 5.7).

5.7.2 Calibration of Maintain Interpersonal Distance Behaviour and Wall Avoidance Behaviour

Calibration process for maintain interpersonal distance and wall avoidance behaviours is consisted of 625 cases (parameter sets) with 4 parameters (β_{ip}, ρ_{ip}, β_w, and ρ_w) and 5 values per parameter. Values for each parameter were selected after performing many simulation trials. These values are as follows:

β_{ip}= [-0.8 -1 -3 -5 -8]

ρ_{ip}= [-0.5 -1 -3 -5 -8]

β_w= [-0.5 -2 -5 -8 -10]

ρ_w= [-0.2 -1 -3 -5 -7]

β_{ddist} *and* β_{ddir} have been fixed to values obtained from previous *sub-section* (-3 and -0.2 respectively). The total of 625 simulation runs have been performed.

The measures of performance selected for this part of the calibration process are egress time, total walking distance, and portion of interpersonal distances below distance threshold. At every time step, the distances between the pedestrian and individuals in front of her (the angle between pedestrian's movement direction and the position of other individuals is between 0° to 90°) have been calculated and finally the interpersonal distance of each pedestrian is defined as the minimum of all these distances. The portion of interpersonal distances below distance threshold

is the percentage of all calculated interpersonal distances during the whole simulation time that are below 40 centimetres.

The egress time and flow obtained for 25 combinations of parameters (out of 625 combinations) are presented in Table 5.8. Those parameter sets with acceptable egress time are selected as candidate parameter sets.

Table 5.8: Sample of simulated egress time for different combinations of parameters

β_{ip}	ρ_{ip}	β_w	ρ_w	Egress time (*sec*)
-0.8	-0.5	-0.5	-0.2	55
-0.8	-1	-0.5	-1	63.5
-0.8	-3	-2	-7	57
-0.8	-5	-8	-1	47.5
-0.8	-8	-8	-3	49
-1	-0.5	-5	-1	65.5
-1	-1	-5	-1	55.5
-1	-3	-8	-3	55
-1	-5	-0.5	-7	58.5
-1	-8	-5	-0.2	45.5
-3	-0.5	-2	-1	59
-3	-1	-10	-1	43.5
-3	-3	-2	-7	60.5
-3	-5	-0.5	-5	59
-3	-8	-0.5	-0.2	63.5
-5	-0.5	-0.5	-7	65.5
-5	-1	-5	-0.2	54
-5	-3	-2	-1	66
-5	-5	-0.5	-3	58.5
-5	-8	-10	-5	49.5
-8	-0.5	-10	-5	51.5
-8	-1	-5	-1	64
-8	-3	-8	-5	67
-8	-5	-0.5	-5	58
-8	-8	-10	-1	56.5

Total walking distance and the percentage of interpersonal distances below threshold have been measured for all candidate sets. The results for some of the candidate sets are shown in Table 5.9. The best parameter set is the one that results in smaller values of total walking distance and percentage of interpersonal distances below threshold, and has been selected based on these two criteria.

Table 5.9: Evaluation of candidate parameter sets based on walking distance and interpersonal distance

β_{ip}	ρ_{ip}	β_w	ρ_w	Egress time (*sec*)	Total walking distance (*m*)	Interpersonal distances below Dth (%)
-0.8	-0.5	-10	-5	51.5	855.53	40.61
-0.8	-1	-10	-5	58.5	902.92	44.75
-0.8	-3	-0.5	-0.2	51.5	834.23	42.75
-0.8	-5	-10	-1	51.5	897.13	52.73
-0.8	-8	-5	-7	52	830.33	58.20
-1	-0.5	-10	-5	51	908.67	42.46
-1	-1	-10	-1	50	854.71	42.17
-1	-3	-8	-3	55	931.39	53.09
-1	-5	-5	-5	53	839.39	42.73
-1	-8	-2	-5	51	839.03	39.63
-3	-0.5	-8	-1	49.5	801.91	38.01
-3	-1	-2	-1	50.5	806.62	40.69
-3	-3	-10	-1	56	791.88	35.23
-3	-5	-0.5	-1	53.5	802.52	32.02
-3	-8	-2	-5	53	784.78	26.11
-5	-0.5	-2	-1	51	842.46	41.34
-5	-1	-2	-0.2	55.5	873.83	38.23
-5	-3	-0.5	-1	49.5	806.96	26.67
-5	-5	-8	-5	53	927.10	33.45
-5	-8	-2	-3	52.5	798.35	24.56
-5	-8	-10	-5	49.5	744.73	18.99
-8	-0.5	-0.5	-3	55	934.64	45.40
-8	-1	-2	-1	56.5	886.45	33.37
-8	-5	-0.5	-5	58	848.78	24.06
-8	-8	-5	-1	57.5	855.29	24.87

The parameter set with β_{ip}, ρ_{ip}, β_w, and ρ_w respectively equal to -5, -8, -10, and -5 has been chosen as the best parameter set based on its walking distance and interpersonal distances (shown in purple in Table 5.9). The total walking distance and the percentage of interpersonal distances below threshold obtained for this combination of parameters are minimum.

The reported values of the percentage of interpersonal distances below threshold in Table 5.9 imply that the current specification for maintain interpersonal distance behaviour can only mitigate the problem present in next step model regarding the distances between pedestrians. The main reason is that the decision-maker always predicts the next step position of nearby individuals with the assumption of constant velocity. However, this is not always the case and individuals may find other alternatives (rather than cell 17) more attractive in some situations. As a result, the decision-maker may avoid the predicted position of another individual nearby which is not really going to be occupied by that individual and instead may choose an alternative which makes her distance to the actually selected position of the individual closer than distance threshold (or in some cases they may even overlap). Therefore, in the current developed model, the goal is to decrease the portion of interpersonal distances that are below threshold as much as possible.

The simulation model was run with no interpersonal distance behaviour and once with no wall avoidance behaviour. The results are reported in Table 5.10.

Table 5.10: Evaluation of two specific parameter sets (no maintain interpersonal distance behaviour and no wall avoidance behaviour)

β_{ip}	ρ_{ip}	β_w	ρ_w	Egress time (*sec*)	Total walking distance (*m*)	Interpersonal distances below Dth (%)
0	0	-5	-5	6.5	627.82	79.61
-5	-8	0	0	74.5	876.60	31.65

When there is no interpersonal behaviour, the egress time and the percentage of interpersonal distance below threshold are extremely low and extremely high respectively. The low value of walking distance is due to the fact that pedestrians do not maintain any distance with each other and freely overlap and walk through each other to reach their destination, which results in very smooth tracks. When there is no wall avoidance behaviour, the egress time is not within the acceptable range and the resulted total walking distance and the percentage of interpersonal distances below threshold are higher than what was obtained based on the best selected parameter set. These results clearly show the dominance of the developed model over next step model in reproducing behaviours of pedestrians walking through bottlenecks in high density situations.

5.7.3 Estimation Results

The estimated values for the 6 parameters β_{ddist}, β_{ddir}, β_{ip}, ρ_{ip}, β_w, and ρ_w are reported in Table 5.11.

Table 5.11: Estimated values for the developed model

Parameter	Estimated value	Parameter	Estimated value
β_{ddir}	-0.2	ρ_{ip}	-8
β_{ddist}	-3	β_w	-10
β_{ip}	-5	ρ_w	-5

Figure 5.14 illustrates the histogram of simulated selected choices for the selected parameter set. Expectedly, alternative 17 is the prevalent choice. Figure 5.15 shows the histogram of interpersonal distances. Figure 5.16 displays the trajectories of all 100 pedestrians for the selected parameter set. Compared to resulted trajectories from next step model (Figure 5.7), these trajectories are unsmooth with more direction changes which are in fact more consistent with the behaviour of pedestrians in high density situations.

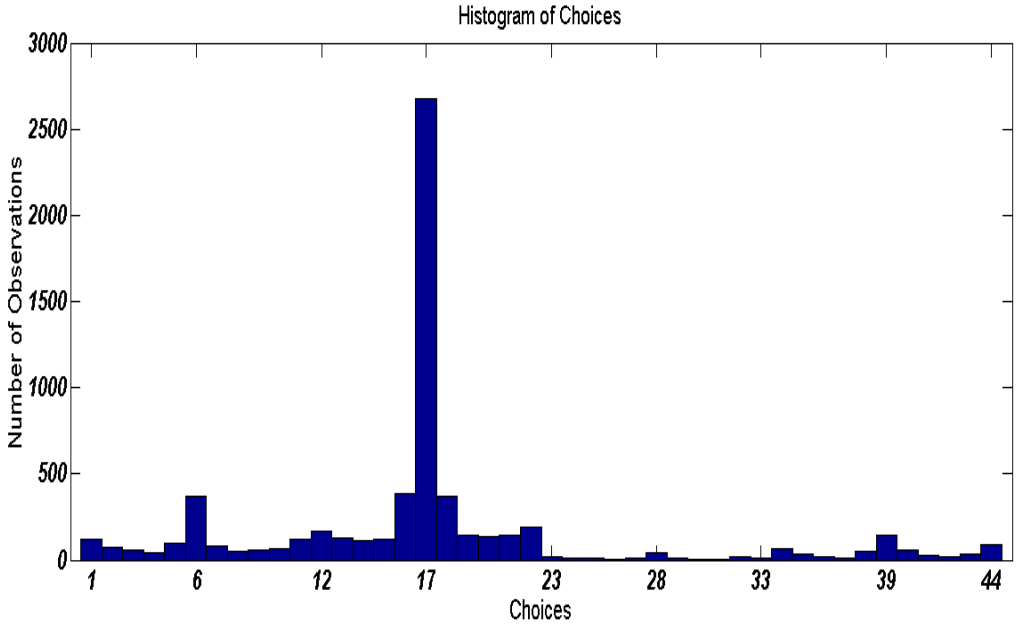


Figure 5.14: Histogram of obtained choices based on developed model

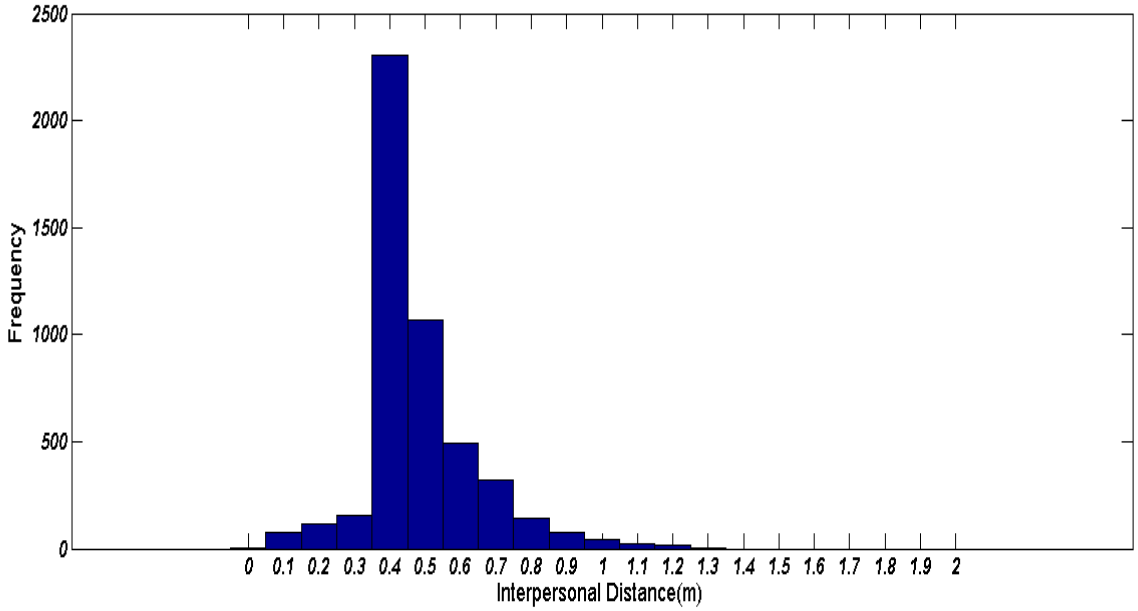


Figure 5.15: Frequency of interpersonal distances based on developed model

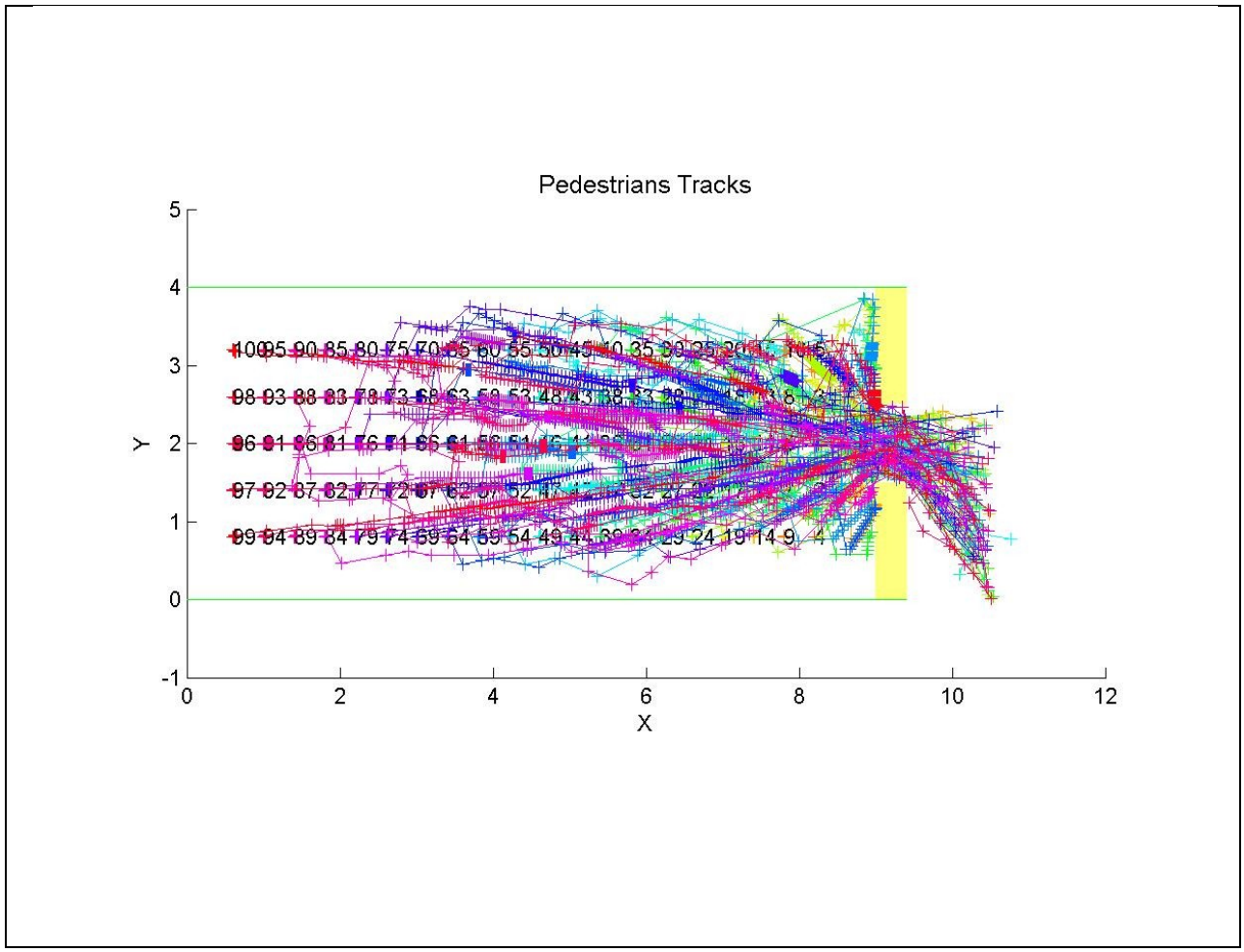


Figure 5.16: Trajectories of all pedestrians based on developed model (axes units are metre)

5.8 Model Validation

Validation is an essential part of model development. Validation process may reveal the requirement for adjustment in calibration process. This section aims to reproduce the field-measured traffic conditions in order to validate the developed model. Two different scenarios have been considered for validation purpose. First, the experiments (except the one used for calibration) conducted by Kretz *et al.* (2006) in bottleneck situations have been simulated using the developed model and the results are compared. Second, the developed model has been

validated based on O-train data described in section 5.2. Both of the validation scenarios are explained in the following.

5.8.1 Duisburg Validation

The developed model has been validated based on experiments conducted by Kretz *et al.* (2006) except for the one used for calibration. The differences between experiments are the bottleneck widths and the number of participants. Bottlenecks of 40, 50, 60, 70, 80, 90, 100, 120, 140, and 160 centimetres wide with 80 and 100 participants have been used for this purpose. The simulated egress time (flow through bottleneck) and average flow are reported in Table 5.12 and are compared with the obtained results from Tobias experiments. Average flow is the average number of pedestrians passing the bottleneck every second during the simulation.

Table 5.12: Simulated egress time and average flow for different bottleneck widths

Bottleneck width (cm)	Number of participants	Egress time (sec)	Average flow (ped/s)
40	80	82	1.03
50	80	73	1.1
60	80	56.5	1.48
70	80	56.5	1.42
80	80	51	1.56
90	80	45	1.73
100	80	44	1.87
120	80	42.5	1.9
140	80	40	1.97
160	80	43	1.73
40	100	104.5	1
50	100	95	1.067
60	100	80	1.31
70	100	66.5	1.51
80	100	60	1.65
90	100	55.5	1.74
120	100	51.5	1.85
140	100	49	2.02
160	100	55.5	1.82

Results in Table 5.12 show the good performance of the model. The fact that the egress time generally decreases (or average flow increases) as the bottleneck width increases reveals the robustness of the model. Overall, the simulated values in the table are consistent with the reported results in tables 5.3 and 5.4. Although some of the obtained egress times or average flows are not within the range of values reported by Kretz *et al.* (2006), they are close to what are expected. It is worth noting that the controlled nature of experiments performed by Kretz *et al.* may affect the outcomes and make the results slightly different from what are expected from natural walking situations. It was observed that for the bottlenecks of higher width (140 or160 cm wide) sometimes the egress time grows (while average flow decreases) compared to the bottlenecks of narrower width. This result is also compatible with the results in table 5.3 and 5.4.

Limitations Associated with the Developed Model

Validation process revealed some of the weak points associated with the developed model, which have to be further investigated.

First, the current free flow term which captures the acceleration behaviour of pedestrians does not function properly in high density situations. In the developed model the decision-maker is provided with the possibility to choose the alternatives with very low speeds. These alternatives are required for high density circumstances when most of the space in front of the pedestrian is blocked with other individuals. However, when pedestrian's space and accordingly the freedom of movement increase, it is expected that the pedestrian accelerate to a higher and a more reasonable speed. However, this is not always the case in the developed model. The current acceleration term captures the tendency of the pedestrian to maintain her current speed and therefore the attractiveness of an alternative with the same low speed is higher than the

attractiveness of an alternative with a higher speed which represents pedestrian's average walking pace. In other words, in some cases even when the area in the front is almost empty the pedestrian who is already walking slowly keeps moving at a very low speed. This behaviour may bring about very long delays. It was observed that sometimes during the simulation no pedestrian passes through the bottleneck for a relatively long time. The graph of cumulative number of pedestrians passing through the door over time for one of these cases is shown in Figure 5.17. During the time steps of 159 to 185 (13 sec) nobody passes the bottleneck.

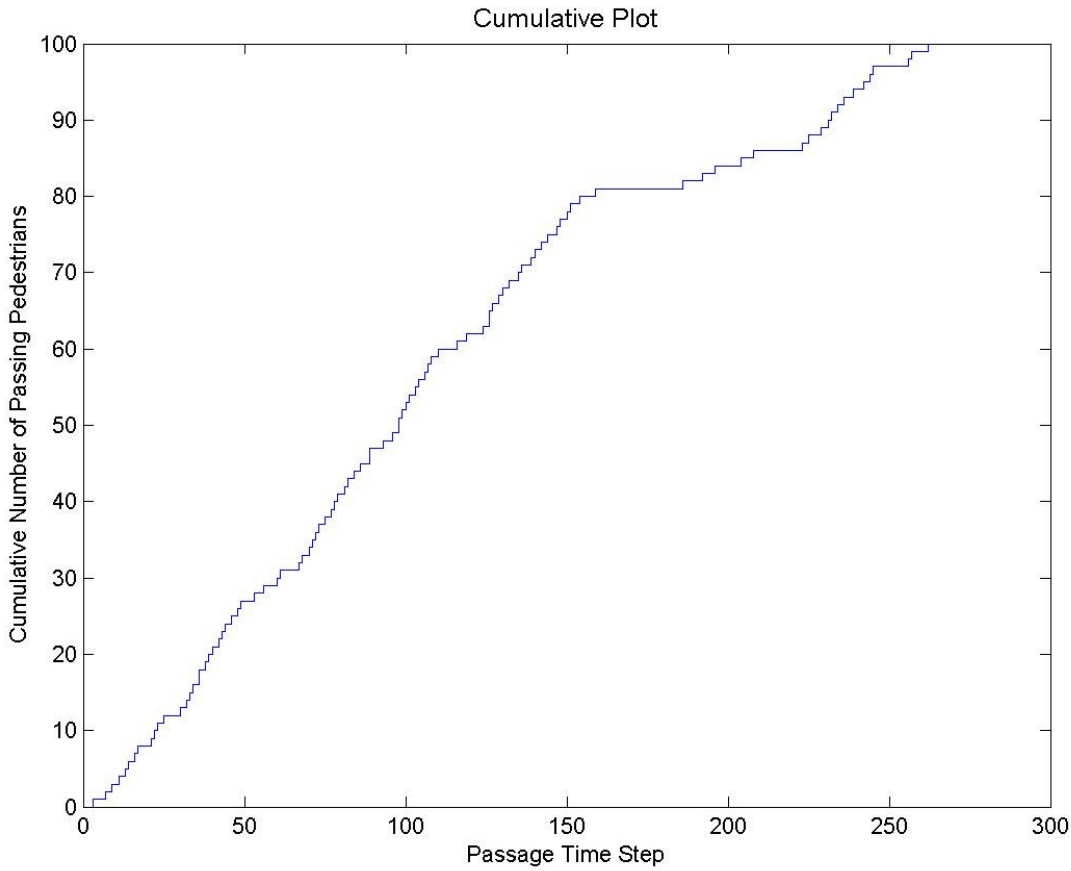


Figure 5.17: Sample of cumulative number of pedestrians passing through bottleneck

It was also figured out that the long delays may have another reason. As mentioned earlier, it was assumed that in case pedestrians block each others' way, the right of way is given to the pedestrian who is closer to the destination (all the pedestrians have similar intermediate destinations at the centre of the bottleneck). However, a more intelligent and sophisticated way of giving way may be required. It was observed that in some cases the closest pedestrian to the destination has the moving direction opposite to other blocked pedestrians. Therefore, giving way to the pedestrian who has to pass through other pedestrians cannot be an efficient way of resolving the blocking problem and brings about delays until finally this pedestrian finds her way among other pedestrians in coming time steps.

Considering the aforementioned limitations in the developed model, if nobody passes through the bottleneck for more than 5 seconds, that duration of time is removed from the calculation of egress time and the average flow of pedestrians in order for the results to be comparable with Kretz *et al.*'s experiments.

5.8.2 O-train Validation

Pedestrians' movements when leaving a train have also been simulated using the developed model for the validation purpose and the results are compared with what was obtained from o-train data collection as reported in section 5.2. Out of 6 conducted observations, the one corresponding to the purple column of table 5.5 has been used for this purpose.

Simulation Setup

The area inside O-train has been built in the simulation tool. This area is a rectangular of approximately 40 by 2.7 metres. Walls have been defined all around the sitting area of

passengers (*i.e.,* chairs) and other obstacles. 215 pedestrians are loaded to the system. Their initial positions have been selected in a way that they maintain at least 40 centimetres distance from nearby pedestrians. A very slow initial speed (0.01 *m/s*) is assigned to all pedestrians. All pedestrians have a final destination which is set outside the train. Also, each pedestrian has 1, 2, or 3 intermediate destinations inside the train. The intermediate destinations depend on where the pedestrian's initial position is and from which door she leaves (A, B, or C). Each door has the width of 1.3 metres and is defined by an opening in the simulation tool. Figure 5.18 shows the positions of pedestrians in a random time step.

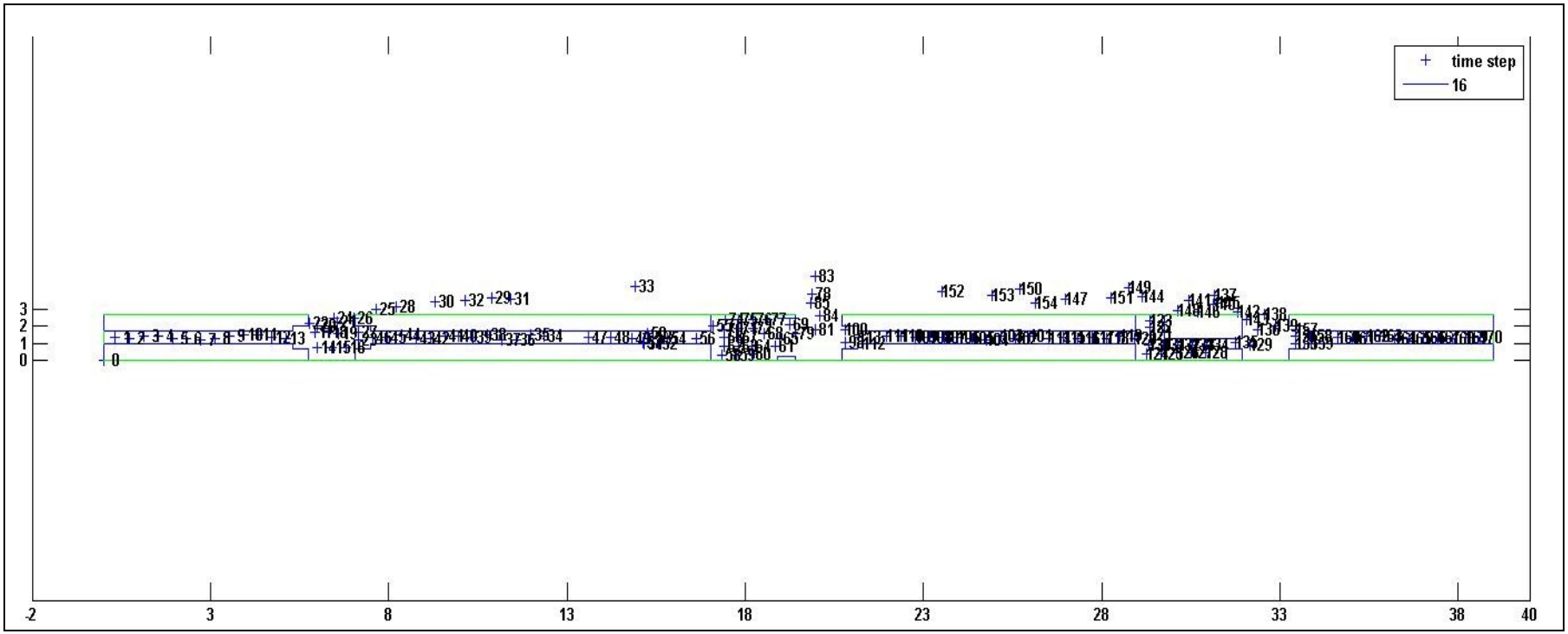


Figure 5.18: A simulated frame of pedestrians' movements when leaving the o-traine (axes units are metre)

Validation Results

The simulated egress time and the simulated average flow for each 5 seconds are reported in Table 5.13.

Table 5.13: Simulated egress time and average flow for each door of the train

Train Door	Number of passengers	Egress time (sec)	Average flow (ped/5s.m)
A	68	55	4.85
B	80	66	4.82
C	67	53.5	5.04

The values of egress time and average flow in Table 5.13 are compatible with the results in Table 5.5 and what was reported in section 5.2. The egress times reported in Table 5.5 were obtained from random observations of the unloading process of the train and hence, it is not expected that the simulation generate the exact same values. The observed train was unloading passengers on campus station during morning time. Obviously, most of the passengers are students who are in a rush to attend their classes on time and this may be the reason for slightly lower egress times compared to what has been obtained from the simulation. The simulated egress time decreases as the number of passengers decreases which is also consistent with O-train dataset and implies the robustness of the model.

Figure 5.19 shows the simulated number of pedestrians passing a unit of length (1 metre) of the train's doors every 5 seconds. The range of flows obtained from simulation is consistent with the observed real data (Figure 5.5).

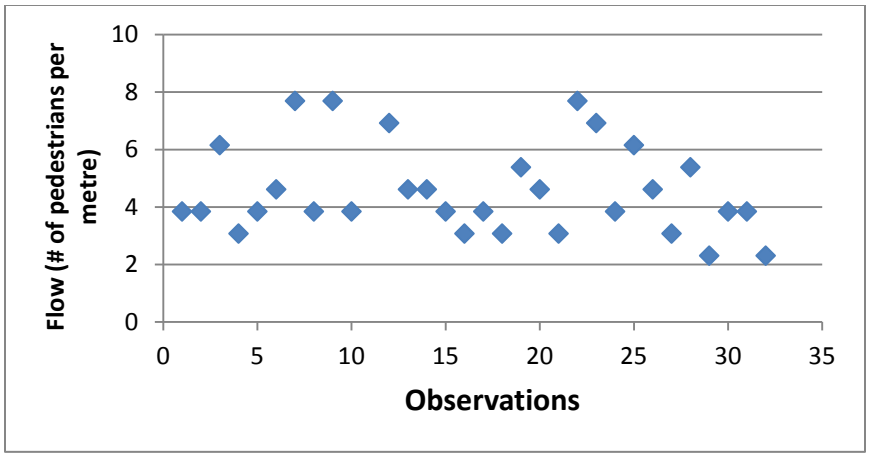


Figure 5.19: Simulated pedestrian flow every 5 seconds (O-train)

The histogram of simulated interpersonal distances is illustrated in Figure 5.20; 15.08 % of interpersonal distances are below the distance threshold (0.4 metre).

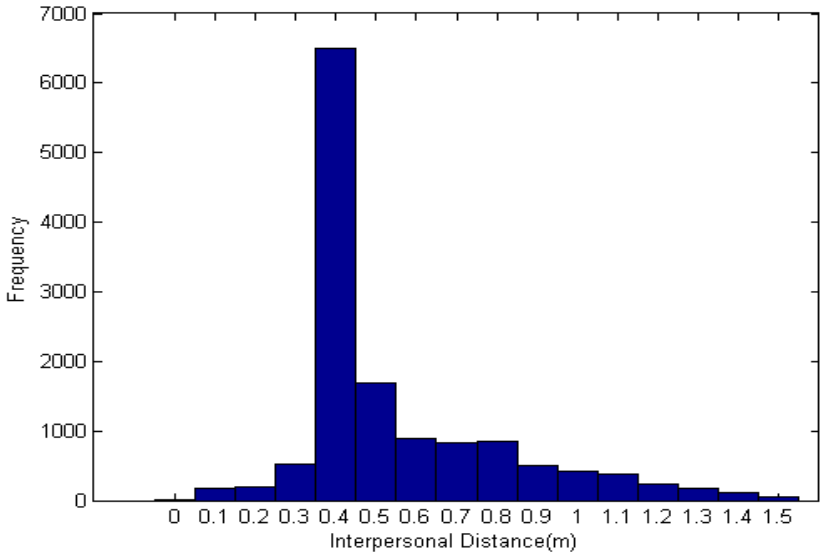
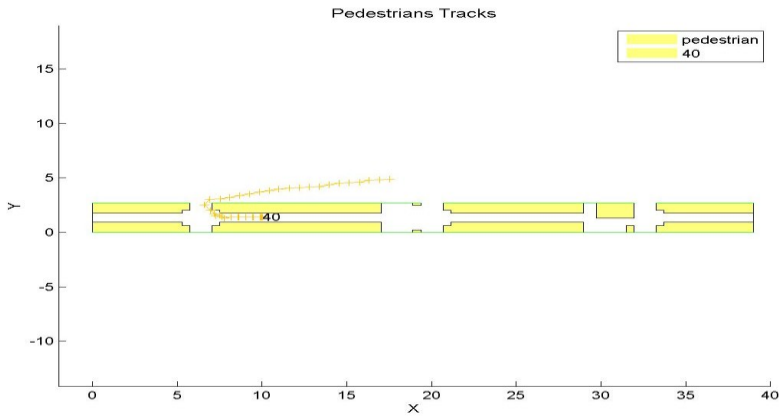


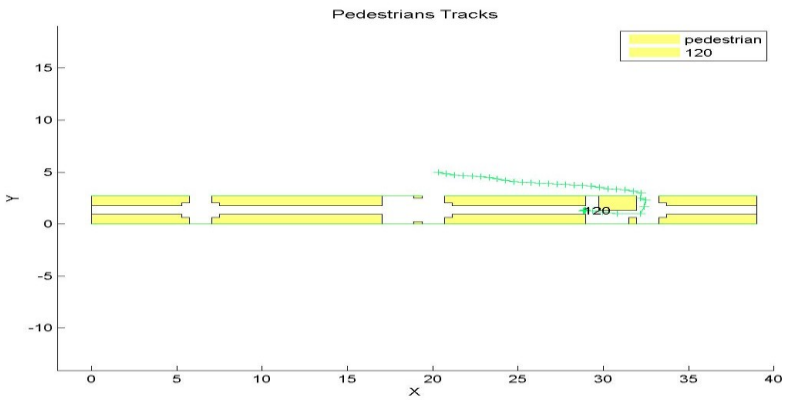
Figure 5.20: Histogram of simulated interpersonal distances (O-train)

Figure 5.21 shows the sample trajectories of two pedestrians as well as the trajectories of all pedestrians resulted from the simulation. These trajectories confirm the good performance of the model as well.

a)



b)



c)

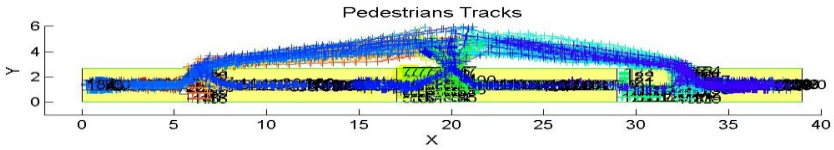


Figure 5.21: Simulated trajectories of (a) and (b) a single pedestrian, (c) all pedestrians exiting the O-train (axes units are metre)

5.9 Summary

It was found out that next step model is not able to properly model pedestrians' behaviour when walking through bottlenecks in high density situations and has its own limitations. In this chapter, next step model has been developed in order to generate a walking behaviour model which is consistent with the discussed situation. For this purpose, two behaviours have been added to the behavioural patterns of next step model, namely, interpersonal distance behaviour and wall avoidance behaviour and also some revisions have been made to the model regarding toward destination behaviour. Moreover, a new row of alternatives has been added to the choice set to capture the tendency of decision-maker to decelerate to a very low speed in high density situations. The calibration and validation results indicate the acceptable performance of the developed model and show its superiority compared to next step model. However, the validation process also has revealed the problems associated with the developed model which should be further investigated.

Chapter six: Conclusion and Future Work

6.1 Research Summary

In general, study of pedestrian walking behaviour using discrete choice models (DCM) are of interest in this research. Discrete choice models concentrate directly on behavioural aspects of pedestrian movement and are capable to capture different patterns of correlation between alternatives as well as the effect of different characteristics of individuals on decision making. These types of models are developed to be calibrated and validated on real data which makes them more realistic representation of pedestrian behaviour.

In this study, group behaviour and pedestrian movements at bottlenecks have been investigated using DCM framework. Study of group behaviour is essential in providing a realistic representation of walking behaviour. Pedestrians belonging to the same group maintain the group unity while walking towards their destinations. In the first part of this thesis, group behaviour has been modeled based on the tendency of group members to stay together (keep short distance with centre of the group) and to match their movement direction with average moving direction of their group-mates. The model has been calibrated and validated based on the real pedestrian data which was collected during a large firework event at Vancouver, Canada (Ismail *et al.* 2009). The majority of pedestrians were walking in groups in this dataset. Groups of 2 to 5 pedestrians have been observed in the crowd. Likelihood ratio test showed that incorporating the effect of

grouping relationships into the walking behaviour model significantly improves its fit to the dataset compared to the model without group behaviour.

The estimation results indicate the significance of this behaviour in modelling pedestrian walking behaviour. The estimated parameters associated with this behaviour are significant and their signs are compatible with behavioural expectations. The validation results also confirm the acceptable performance and stability of the model. Modelling group behaviour based on discrete choice framework is a novel contribution in this research.

Investigation of pedestrian behaviour in critical situations such as high density conditions is of vital importance from different aspects, especially in terms of pedestrian safety. In the second part of the thesis, a simulation tool has been developed in order to model the behaviour of individuals at bottlenecks and under high density situations. The research conducted in this thesis shows that the discrete choice model proposed by Robin *et al.* (2009) which is known as next step model is not able to accurately reproduce pedestrian movements at mentioned situations. Therefore, next step model is developed and revised to represent realistic behaviour of pedestrians. Behavioural patterns have been extended by adding contributions to the utility function capturing wall avoidance behaviour and maintenance of interpersonal distance behaviour. Some revisions regarding the model specification and choice set structure have also been made.

For the sake of calibration, different parameter sets have been considered and the best set has been selected based on different performance measures. The developed model has been calibrated and validated based on the data extracted from video recordings of controlled experiments conducted by Kretz *et al.* (2006). In total 20 different experiments were conducted

for different bottleneck widths and different number of participants. The experiment corresponding to the bottleneck width of 1 metre and 100 participants has been selected for calibration. Validation has been performed based on all 19 other experiments. Moreover, the performance of the model in simulating the behaviour of pedestrians leaving a train has been examined and compared to the data collected at O-Train station at Ottawa, Canada for validation purpose. The calibration and validation process show the superiority of the developed model compared to next step model in reproducing behaviour of high density crowd at bottlenecks. However, some shortcomings associated with the developed model have been revealed in validation process which have to be overcome by further developing the model. Study of crowd dynamics at bottlenecks in a discrete choice framework is also an innovative contribution of this research.

It should be noted that pedestrians do not behave the same under different circumstances. Therefore, the discrete choice walking models proposed in this study can only be generalized to similar contexts.

6.2 Future Work

The validation process of group behaviour has been restricted to the choices selected by decision-makers. The model was not validated based on macroscopic pedestrian flow characteristics such as flow and density or based on self organized patterns forming in pedestrian crowd such as dynamic lanes. As a future work, it is intended to implement the model in the simulation tool and validate it more thoroughly at both microscopic and macroscopic levels. Moreover, it is aimed to validate the model based on the dataset other than the one used for calibration purpose.

The groups of different size can be observed in pedestrian crowds. The size of the group may influence pedestrian walking behaviour. This effect has not been considered in this study. A future development would be to study the influence of different group size on walking behaviour of pedestrians. Groups of different size may form different shapes in the crowd. Also, in larger groups as opposed to groups of small size pedestrians may tend to match their walking behaviour with close group-mates rather than all the pedestrians in the group. These potential patterns in behaviour of pedestrians need to be further investigated.

The utility term capturing the free flow acceleration behaviour has to be improved. The assumption that individuals tend to maintain their current speed may not be rational in all circumstances. Speed of individuals besides other factors depends on density of the crowd. When a pedestrian is walking with a very low speed in unimpeded situations, it is more realistic to assume that she is more prone to select an accelerated alternative rather than the one with constant speed. Therefore, associating acceleration alternatives with negative utility may not be reasonable in these cases.

The maintenance of interpersonal distance behaviour has been introduced in chapter 5 to fix the problem of having interpersonal distances below threshold in high density situations. However, it turns out that this behaviour can only partially resolve the mentioned problem. A further development of the model would be to improve this behaviour in a way that no two pedestrians overlap or walk through each other. It should be noted that in contrast to what is assumed in time-based simulations such as the simulation tool developed in this study, in reality all pedestrians do not update their positions at the exact same time instant. Hence, one solution to produce acceptable interpersonal distances is to develop an event-based simulation and update pedestrians' positions one by one.

Giving the right of way to a single pedestrian when a group of pedestrians are blocking each others' way has been considered as a strategy to fix the freezing problem in simulation tool. The validation results in chapter 5 showed that giving the right of way to the pedestrian who is closer to the destination cannot always resolve the problem and a more intelligent method for giving way is required. This issue should be further investigated.

One area of development is to incorporate the effect of pedestrians' characteristics into walking behaviour model. Factors such as age, gender, and height affect how pedestrians move. However, extracting these attributes from recordings of natural pedestrian movements is not an easy task and can even be impossible in some cases. Conducting controlled experiments is one way to capture these characteristics in pedestrian crowd.

Another potential future work is to develop the proposed discrete choice model to capture the behaviour of pedestrians in panic situations such as emergency evacuation scenarios. The factors influencing pedestrian behaviour in these situations may differ from those of normal situations. For instance, the leader-follower behaviour may play a stronger role in decision making behaviour of individuals while the collision avoidance behaviour may have a weaker effect on utility of alternatives compared to normal situations.

References

1. Antonini, G., 2005. A discrete choice modeling framework for pedestrian walking behavior with application to human tracking in video sequences. *Faculty of Science and Technology in Engineering*, 3382.
2. Asano, M., Iryo, T. & Kuwahara, M., 2010. Microscopic pedestrian simulation model combined with a tactical model for route choice behaviour. *Transportation Research Part C: Emerging Technologies*, 18(6), pp.842–855.
3. Aveni, A., 1977. The not-so-lonely crowd: Friendship groups in collective behavior. *Sociometry*, 40(1), pp.96–99.
4. Aycin, M. & Benekohal, R., 1998. Linear acceleration car-following model development and validation. *Transportation Research Record*, (98), pp.10–19.
5. Bailey, K., Hartnett, J. & Gibson, F.J., 1972. Implied threat and the territorial factor in personal space. *Psychological reports*, 30(1), pp.263–270.
6. Bandini, S. et al., 2012. Crowd and Pedestrian Dynamics: Empirical Investigation and Simulation. *measuringbehavior.org*, 2012, pp.308–311.
7. Bando, M. et al., 1995. Dynamical model of traffic congestion and numerical simulation. *Physical Review E*, 51(2), pp.1035–1042.
8. Ben-Akiva, M & Bierlaire, M, 1999. Discrete choice methods and their applications to short-term travel decisions. In R. Hall, ed. *Handbook of Transportation Science*. Kluwer, pp. 7–38.
9. Ben-Akiva, M & Bierlaire, M, 2003. Discrete choice models with applications to departure time and route choice. In *Handbook of transportation science*. Springer US, pp. 7–37.

10. Ben-Akiva, M & Lerman, S.R., 1985. *Discrete Choice Analysis: Theory and Applications to Travel Demand* M. L. Manheim, ed., MIT Press.

11. Ben-Akiva, ME, 1974. Structure of passenger travel demand models. *Transportation Research Record*.

12. Van Den Berg, J. et al., 2008. Interactive navigation of multiple agents in crowded environments. *Proceedings of the 2008 symposium on Interactive 3D graphics and games SI3D 08*, 1(212), p.139.

13. Van Den Berg, J. & Manocha, D., 2008. Reciprocal Velocity Obstacles for real-time multi-agent navigation. *2008 IEEE International Conference on Robotics and Automation*, pp.1928–1935.

14. Bierlaire, M., 2003. BIOGEME: A free package for the estimation of discrete choice models. In *Proceedings of the 3rd Swiss Transportation Research Conference*.

15. Bierlaire, M. & Robin, T., 2009. Pedestrian choices. In H. Timmermans, ed. *Pedestrian Behavior: Models, Data Collection and Applications*. Emerald Group, pp. 1–26.

16. Blue, V. & Adler, J., 2000. Modelling four directional pedestrian movements. *Transportation Research Board, 79th Annual Meeting*, 1710(1), pp.20–27.

17. Blue, Victor & Adler, Jeffrey, 1998. Bi-directional emergent fundamental pedestrian flows from cellular automata microsimulation A Ceder, ed. *Transportation Research Record*, 1644(-1), pp.29–36.

18. Blue, Victor & Adler, Jeffrey, 2001. Cellular automata microsimulation for modeling bi-directional pedestrian walkways. *Transportation Research Part B: Methodological*, 35(3), pp.293–312.

19. Blue, Victor & Adler, Jeffrey, 1999. Cellular Automata Microsimulation of Bidirectional Pedestrian Flows. *Transportation Research Record*, 1678(1), pp.135–141.

20. Borgers, A. & Timmermans., H., 1986. City centre entry points, store location patterns and pedestrian route choice behaviour: A microlevel simulation model. *Socio-economic planning sciences*, 20(I), pp.25–31.

21. Borghese, N.A., Bianchi, L. & Lacquaniti, F., 1996. Kinematic determinants of human locomotion. *The Journal of Physiology*, 494 (Pt 3(3), pp.863–79.

22. Börsch-Supan, A., 1990. On the Compatibility of Nested Logit Models with Utility Maximization. *Journal of Econometrics*, 43(3), pp.373–388.

23. Burstedde, C et al., 2001. Simulation of pedestrian dynamics using a 2-dimensional cellular automaton. *Physica A: Statistical Mechanics and its Applications*, 295(3-4), p.22.

24. Burstedde, Carsten et al., 2001. Cellular Automaton Approach to Pedestrian Dynamics - Applications M Schreckenberg & S. D. Sharma, eds. *condmat0112119*, p.11.

25. Camillen, F. et al., 2009. Multi agent simulation of pedestrian behavior in closed spatial environments. In *Science and Technology for Humanity TICSTH 2009 IEEE Toronto International Conference*. pp. 375–380.

26. Chandler, R.E., Herman, R. & Montroll, E.W., 1958. Traffic Dynamics: Studies in Car Following. *Operations Research*, 6(2), pp.165–184.

27. Cheung, C. & Lam, W., 1998. Pedestrian route choices between escalator and stairway in MTR stations. *Journal of transportation engineering*, 2(May 2012), pp.277–285.

28. Chraibi, M. et al., 2009. Quantitative Description of Pedestrian Dynamics with a Force based Model. *2009 IEEEWICACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, 3, pp.583–586.

29. Coleman, J.S. & James, J., 1961. The Equilibrium Size Distribution of Freely-Forming Groups. *Sociometry*, 24(1), pp.36–45.

30. Crowe, A. et al., 1996. The influence of walking speed on parameters of gait symmetry determined from ground reaction forces. *Human Movement Science*, 15(3), pp.347–367.

31. Daamen, W., 2004. *Modelling passenger flows in public transport facilities*.

32. Daamen, W., 2005. Passenger Route Choice concerning Level Changes in Railway Stations. , 2005, pp.1–18.

33. Daamen, W. & Hoogendoorn, S., 2003. Experimental Research of Pedestrian Walking Behavior. *Transportation Research Record*, 1828(1), pp.20–30.

34. Daly, P.N., McGrath, F. & Annesley, T.J., 1991. Pedestrian speed/flow relationships for underground stations. , 32(2), pp.75–78.

35. Dosey, M. a & Meisels, M., 1969. Personal space and self-protection. *Journal of personality and social psychology*, 11(2), pp.93–7.

36. Durupinar, F. et al., 2008. Creating Crowd Variation with the OCEAN Personality Model. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 3*. International Foundation for Autonomous Agents and Multiagent Systems, pp. 1217–1220.

37. Dzubiella, J., Hoffmann, G.P. & Löwen, H., 2002. Lane formation in colloidal mixtures driven by an external field. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 65(2 Pt 1), p.021402.

38. Feurtey, F., 2000. *Simulating the collision avoidance behavior of pedestrians*.

39. Fiorini, P. & Shiller, Z., 1998. Motion Planning in Dynamic Environments Using Velocity Obstacles. *The International Journal of Robotics Research*, 17(7), pp.760–772.

40. Forsyth, D.R., 2010. *Group Dynamics*, Wadsworth Publishing.

41. Foudil, C. & Noureddine, D., 2006. Collision avoidance in crowd simulation with priority rules. *European Journal of Scientific Research*, 15(1), pp.6–17.

42. Frankel, A. & Barrett, J., 1971. Variations in personal space as a function of authoritarianism, self-esteem, and racial characteristics of a stimulus situation. *Journal of Consulting and Clinical Psychology1*, 37(1), pp.95–98.

43. Fridman, N. & Kaminka, G., Towards a cognitive model of crowd behavior based on social comparison theory. In *Proceedings of the National Conference on Artificial Intelligence Vol. 22. No. 1.* Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press;, pp. 731–737.

44. Fruin, J.J., 1971. *Pedestrian planning and design*,

45. Fruin, J.J., 2002. The causes and prevention of crowd disasters. , pp.1–10.

46. Fukui, M.. & Ishibashi, Y., 1999. Self-organized phase transitions in cellular automaton models for pedestrians. *Journal of the Physical Society of Japan*, 68(8), pp.2861–2863.

47. Gazis, D.C., Herman, R. & Potts, R.B., 1959. Car-Following Theory of Steady-State Traffic Flow. *Operations Research*, 7(4), pp.499–505.

48. Ghasemaghaee, N. & Oren, T., 2007. Cognitive complexity and dynamic personality in agent simulation. *Computers in Human Behavior*, 23(6), pp.2983–2997.

49. Goffman, E., 1971. *Relations in Public: Microstudies of the Public Order*, New York: Basic Books.

50. Guy, S et al., 2009. ClearPath : Highly Parallel Collision Avoidance for Multi-Agent Simulation. In D. Fellner & S. Spencer, eds. *Computing*. ACM, pp. 177–187.

51. Hall, E., 1966. The hidden dimension. *Chemistry &*

52. Hartnett, J., Bailey, K. & Gibson, F.J., 1970. Personal space as influenced by sex and type of movement. *The Journal of Psychology: Interdisciplinary and Applied*, 76(2), pp.139–144.

53. Hayduk, L.A., 1983. Personal space: Where we now stand. *Psychological Bulletin*, 94(2), pp.293–335.

54. Hediyeh, H., 2012. *Investigation of microscopic pedestrian walking behavior*. The university of British Colombia.

55. Heïgeas, L. et al., 2010. A physically-based particle model of emergent crowd behaviors. *Graphicon*, 2, pp.5–10.

56. Helbing, D et al., 2005. Self-organized pedestrian crowd dynamics: Experiments, simulations, and design solutions. *Transportation Science*, 39(1), pp.1–24.

57. Helbing, D, Farkas, I. & Vicsek, T., 2000. Simulating dynamical features of escape panic. *Nature*, 407(6803), pp.487–90.

58. Helbing, D & Molnár, P, 1995. Social force model for pedestrian dynamics. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, 51(5), pp.4282–4286.

59. Helbing, D., 1992. A Fluid Dynamic Model for the Movement of Pedestrians. *Arxiv preprint condmat9805213*, 6, pp.391–415.

60. Helbing, D. & Johansson, A., 2009. *Pedestrian, crowd and evacuation dynamics*, New York: New York.

61. Helbing, Dirk et al., 2001. Self-organizing pedestrian movement. *Environment and Planning B: Planning and Design*, 28(3), pp.361–383.

62. Helbing, Dirk, Johansson, Anders & Al-Abideen, Habib Zein, 2007. Dynamics of crowd disasters: An empirical study. *Physical Review E - Statistical, Nonlinear and Soft Matter Physics*, 75(4 Pt 2), p.046109.

63. Helbing, Dirk, Keltsch, J. & Molnar, P., 1998. Modelling the evolution of human trail systems. *arXiv preprint cond-mat/9805158*, pp.1–11.

64. Helbing, Dirk & Molnar, P., 1998. Self-organization phenomena in pedestrian crowds. *arXiv preprint cond-mat/9806152*, pp.1–10.

65. Helly, W., 1961. Simulation in bottlenecks in single-lane traffic flow. In *Proceedings of the Symposium on the Theory of Traffic Flow*. pp. 207–238.

66. Henderson, L.F., 1974. On the fluid mechanics of human crowd motion. *Transportation Research*, 8(6), pp.509–515.

67. Herman, R. et al., 1958. Traffic dynamics: Analysis of stability in car following. *Operations research*, 7(1), pp.86–106.

68. Hess, S. et al., 2012. A joint model for vehicle type and fuel type choice: evidence from a cross-nested logit study. *Transportation*, 39(3), pp.593–625.

69. Hess, S., 2005. Cross-Nested Logit Modelling of the Combined Choice of Airport, Airline and Access-Mode. *European Transport Conference, 2005*, 44(0), pp.1–32.

70. Himanen, V. & Kulmala, R., 1988. An application of logit models in analysing the behaviour of pedestrians and car drivers on pedestrian crossings. *Accident analysis and prevention*, 20(3), pp.187–197.

71. Hoffmann, N., 2000. The Capacity Restraint Vine: A Powerful Framework for Modelling Individual Travellers Dynamic Decision Making in a Network at Micro Level. In

Proceedings of seminar k of the European transport conference, Homerton college, Cambridge, UK.

72. Hoogendoorn, S, 2003. Pedestrian travel behavior modeling. In *10th International Conference on Travel Behavior Research*. Lucerne, pp. 10–15.
73. Hoogendoorn, S & Bovy, P., 2001. Nonlocal continuous-space microscopic simulation of pedestrian flows with local choice behavior. *Transportation Research Record*, 1776(01), pp.201–210.
74. Hoogendoorn, S & Bovy, P., 2004. Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2), pp.169–190.
75. Hoogendoorn, S, Bovy, P. & Daamen, W., 2002. Microscopic pedestrian wayfinding and dynamics modelling. *Pedestrian and Evacuation Dynamics*, pp.123–154 452.
76. Hoogendoorn, SP, 2004. Walking behavior in bottlenecks and its implications for capacity. *Transportation Research Board*, pp.1–13.
77. Hoogendoorn, SP & Daamen, W., 2005. Pedestrian behavior at bottlenecks. *Transportation Science*, 39(2), pp.147–159.
78. Horowitz, M.J., Duff, D.F. & Stratton, L.., 1964. Body-buffer zone; Exploration of personal space. *Archives of general psychiatry*, 11, pp.651–656.
79. Hughes, R L, 2000. The flow of large crowds of pedestrians. *Mathematics and Computers in Simulation*, 53(4-6), pp.367–370.
80. Hughes, Roger L, 2003. The Flow of Human Crowds. *Annual Review of Fluid Mechanics*, 35(1), pp.169–182.

81. Hui, X. et al., 2007. Pedestrian Walking Speed, Step Size, and Step Frequency from the Perspective of Gender and Age: Case Study in Beijing, China.". *Transportation Research Board 86th Annual Meeting, no. 07-1486*.

82. Ishaque, M.M. & Noland, R.B., 2008. Behavioural Issues in Pedestrian Speed Choice and Street Crossing Behaviour: A Review. *Transport Reviews*, 28(1), pp.61–85.

83. Ismail, K., Sayed, T. & Saunier, N., 2013. A methodology for precise camera calibration for data collection applications in urban traffic scenes. *Canadian Journal of Civil Engineering*, Volume 40(Number 1), pp.57–67(11).

84. Ismail, K., Sayed, Tarek & Saunier, Nicolas, 2009. Automated Collection of Pedestrian Data Using Computer Vision Techniques. *Transportation Research Board Annual Meeting Compendium of Papers*.

85. Jaganathan, S., Clarke, T.L. & Kaup, D.J., 2007. Intelligent Agents- Incorporating Personality into Crowd Simulation. In *InterserviceIndustry Training Simulation and Education Conference2*. pp. 2–9.

86. Ji, Q. & Gao, C., 2006. Simulating crowd evacuation with a leader-follower model. *IJCSES*, 1(4), pp.249–252.

87. Johansson, Anders et al., 2008. From Crowd Dynamics to Crowd Safety: A Video-Based Analysis. *Advances in Complex Systems*, 11(04), pp.497–527.

88. Karamouzas, I. et al., 2009. A predictive collision avoidance model for pedestrian simulation. *Motion in Games*, pp.41–52.

89. Kashani Jou, K., 2011. Pedestrian Areas and Sustainable Development. *World Academy of Science, Engineering and Technology*, pp.483–490.

90. Kingsley Zipf, G., 1949. Human Behavior and the Principle of Least Effort. *Addison-Wesley, Cambody Mus. Am. Arch. and Ethnol.(Harvard Univ.)*, 19, pp.1–125.

91. Kirchner, A., Nishinari, K. & Schadschneider, Andreas, 2002. Friction effects and clogging in a cellular automaton model for pedestrian dynamics. *Physical Review E - Statistical, Nonlinear and Soft Matter Physics*, 67(5 Pt 2), p.056122.

92. Klüpfel, H, 2012. *A cellular automaton model for crowd movement and egress simulation*, GRIN Verlag.

93. Klüpfel, Hubert, Schreckenberg, Michael & Meyer-König.l, T., 2005. Models for crowd Movement and egress Simulation. In *Traffic and Granular Flow'03*. Springer Berlin Heidelberg, pp. 357–372.

94. Komentani, E. & Sasaki, T., 1958. On the stability of traffic flow. *Journal of Operations Research*, 2, pp.11–26.

95. Koppelman, F.S. & Sethi, V., 2005. Incorporating variance and covariance heterogeneity in the Generalized Nested Logit model: an application to modeling long distance travel choice behavior. *Transportation Research Part B: Methodological*, 39(9), pp.825–853.

96. Kretz, Tobias, Grünebohm, A. & Schreckenberg, Michael, 2006. Experimental study of pedestrian flow through a bottleneck. *Journal of Statistical Mechanics: Theory and Experiment*, 2006(10), pp.P10014–P10014.

97. Kuethe, J., 1962. Social schemas. *The Journal of Abnormal and Social Psychology*, 64(1), pp.31–8.

98. Lakoba, T.I., 2005. Modifications of the Helbing-Molnar-Farkas-Vicsek Social Force Model for Pedestrian Evolution. *Simulation*, 81(5), pp.339–352.

99. Lam, W.H.K. & Cheung, C.Y., 2000. Pedestrian speed/flow relationships for walking facilities in Hong Kong. *Transportation Engineering*, 126(4), pp.343–349.

100. Lam, William H K, Lee, J.Y.S. & Cheung, C Y, 2002. A study of the bi-directional pedestrian flow characteristics at Hong Kong signalized crosswalk facilities. *Transportation*, 29(1991), pp.169–192.

101. Lau, M. & Kuffner, J., 2005. Behavior planning for character animation. *Proceedings of the 2005 ACM SIGGRAPHEurographics symposium on Computer animation SCA 05*, (July), p.271.

102. Lee, Y., 2005. *Pedestrian Walking and Choice Behavior on Stairways and Escalators in Public*. Delft University of Technology.

103. Lemp, J.D., Kockelman, K.M. & Damien, P., 2010. The continuous cross-nested logit model: Formulation and application for departure time choice. *Transportation Research Part B: Methodological*, 44(5), pp.646–661.

104. Li, L. et al., 1999. Coordination patterns of walking and running at similar speed and stride frequency. *Human Movement Science*, 18(1), pp.67–85.

105. Li, T.-Y., Jeng, Y.-J. & Chang, S.-I., 2001. Simulating virtual human crowds with a leader-follower model. In *Proceedings Computer Animation 2001 Fourteenth Conference on Computer Animation Cat No01TH8596*. IEEE Comput. Soc, pp. 93–102.

106. Liu, X., Song, Weiguo & Zhang, Jun, 2009. Extraction and quantitative analysis of microscopic evacuation characteristics based on digital image processing. *Physica A: Statistical Mechanics and its Applications*, 388(13), pp.2717–2726.

107. Loscos, C., Marchal, D. & Meyer, A., 2003. Intuitive Crowd Behaviour in Dense Urban Environments using Local Laws. *Proceedings of Theory and Practice of Computer Graphics 2003*, 0, pp.122–129.

108. Ma, J., Song, W, Fang, Z., et al., 2010. Experimental study on microscopic moving characteristics of pedestrians in built corridor based on digital image processing. *Building and Environment*, 45(10), pp.2160–2169.

109. Ma, J., Song, W, Zhang, J, et al., 2010. k-Nearest-Neighbor interaction induced self-organized pedestrian counter flow. *Physica A: Statistical Mechanics and its Applications*, 389(10), pp.2101–2117.

110. Manski, C.F., 1977. The structure of random utility models. *Theory and Decision*, 8(3), pp.229–254.

111. Manual, 2010. *Highway capacity*, Washington,DC.

112. Matsumoto, Y. et al., 1972. (1972). "A Study on Design of Pedestrian Over-Bridges." Transactions of JSCE, 4, 50–51. *Transactions of JSCE*, 4, pp.50–51.

113. McFadden, D. & Train, K., 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics*, 15(5), pp.447–470.

114. McFadden, Daniel, 1981. Econometric Models of Probabilistic Choice. In C. F. Manski & Daniel McFadden, eds. *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, pp. 198–272.

115. McFadden, Daniel, 1978. Modelling the choice of residential location. In A. Karlqvist et al., eds. *Spatial Interaction Theory and Planning Models*. North-Holland, pp. 75–96.

116. Milinskii, A.I. & Predtechenskii, V.M., 1978. *Planning for Foot Traffic Flow in Buildings*, Amerind.

117. Moussaïd, M. et al., 2010. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PloS one*, 5(4), p.e10047.

118. Muramatsu, M., Irie, T. & Nagatani, T., 1999. Jamming transition in pedestrian counter flow. *Physica A*, 267(3-4), pp.487–498.

119. Musse, S & Thalmann, D., 1997. A model of human crowd behavior: Group inter-relationship and collision detection analysis. *Computer Animation and Simulation'97. Springer Vienna*, pp.39–51.

120. Navin, F.P.D. & Wheeler, R.J., 1969. Pedestrian flow characteristics. *Traffic Engineering*, pp.30–36.

121. Newell, G.F., 1961. Nonlinear Effects in the Dynamics of Car Following. *Operations Research*, 9(2), pp.209–229.

122. Newman, R. & Pollack, D., 1973. Proxemics in deviant adolescents. *Journal of Consulting and Clinical Psychology*, 40(1), pp.6–8.

123. O'Flaherty, C.A. & Parkinson, M.H., 1972. Movement on a city centre footway. *Traffic Engineering & Control*, (13), pp.434–438.

124. Osaragi, T., 2004. Modeling of pedestrian behavior and its applications to spatial evaluation. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems 2004 AAMAS 2004*. ACM Press, pp. 834–841.

125. Papadimitriou, E., Yannis, G. & Golias, J., 2009. A critical assessment of pedestrian behaviour models. *Transportation Research Part F: Traffic Psychology and Behaviour*, 12(3), pp.242–255.

126. Paris, S., Pettré, J. & Donikian, S., 2007. Pedestrian Reactive Navigation for Crowd Simulation: a Predictive Approach. *Computer Graphics Forum*, 26(3), pp.665–674.

127. Pauls, J., 1987. Calculating evacuation times for tall buildings. *Fire Safety Journal*, 12(3), pp.213–236.

128. Pelechano, N, Allbeck, J.M. & Badler, N.I., 2007. Controlling individual agents in high-density crowd simulation. In D. Metaxas & J. Popovic, eds. *Proceedings of the 2007 ACM SIGGRAPHEurographics symposium on Computer animation*. Eurographics Association.

129. Perez, G.J. et al., 2002. Streaming, disruptive interference and power-law behavior in the exit dynamics of confined pedestrians. *Physica A: Statistical Mechanics and its Applications*, 312(3-4), pp.609–618.

130. Pipes, L.A., 1953. An Operational Analysis of Traffic Dynamics. *Journal of Applied Physics*, 24(3), pp.274–281.

131. Pushkarev, B. & Zupan, J., 1975. Urban Space for Pedestrians. *MIT PRESS.*

132. Qiu, F., 2010. Modeling dynamic groups for agent-based pedestrian crowd simulations. *Proceedings of the 2010 IEEE/WIC/ACM International*, pp.461–464.

133. Qiu, F. & Hu, X., 2010. Modeling group structures in pedestrian crowd simulation. *Simulation Modelling Practice and Theory*, 18(2), pp.190–205.

134. Reynolds, C., 1987. Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, 21(4), pp.25–34.

135. Reynolds, C.W., 1999. Steering behaviors for autonomous characters. *Game Developers Conference*, 1999(9602), pp.763–782.

136. Robin, T. et al., 2009. Specification, estimation and validation of a pedestrian walking behavior model. *Transportation Research Part B: Methodological*, 43(1), pp.36–56.

137. Saboia, P. & Goldenstein, S., 2012. Crowd simulation: applying mobile grids to the social force model. *The Visual Computer*, 28(10), pp.1039–1048.

138. Sahaleh, A., Bierlaire, Michel & Farooq, B., 2012. Scenario Analysis of Pedestrian Flow in Public Spaces. In *Proceedings of the 12th Swiss Transport Research Conference (STRC)*. pp. 2–4.

139. Sarmady, S., Haron, F. & Talib, A.Z.H., 2009. Modeling Groups of Pedestrians in Least Effort Crowd Movements Using Cellular Automata. *2009 Third Asia International Conference on Modelling & Simulation*, pp.520–525.

140. Schadschneider, A. et al., 2009. Evacuation dynamics: Empirical results, modeling and applications. *Encyclopedia of Complexity and System Science. Springer*, pp.517–550.

141. Schadschneider, Andreas & Seyfried, Armin, 2009. Empirical results for pedestrian dynamics and their implications for cellular automata models. *Pedestrian Behavior: Data Collection and Applications*, pp.27–43.

142. Seyfried, A et al., 2008. Enhanced empirical data for the fundamental diagram and the flow through bottlenecks W. W. F. Klingsch et al., eds. *Review Literature And Arts Of The Americas*, p.12.

143. Seyfried, A et al., 2005. The fundamental diagram of pedestrian movement revisited. *Journal of Statistical Mechanics: Theory and Experiment*, 2005(10), pp.P10002–P10002.

144. Shao, W. & Terzopoulos, D., 2007. Autonomous pedestrians K. Anjyo & P. Faloutsos, eds. *Graphical Models*, 69(5-6), pp.246–274.

145. Shiwakoti, N. & Nakatsuji, T., 2005. Pedestrian simulations using microscopic and macroscopic model. *Infrastructure Planning*, pp.1–4.

146. Sommer, R., 1969. *Personal Space: The Behavioral Basis of Design*, Prentice-Hall.

147. Spellucci, P., 1998a. A new technique for inconsistent QP problems in the SQP method. *Mathematical Methods of Operations Research*, 47(3), pp.355–400.

148. Spellucci, P., 1998b. An SQP method for general nonlinear programs using only equality constrained subproblems. *Mathematical Programming*, 82(3), pp.413–448.

149. Spellucci, P., 1993. *DONLP2 users guide*. Technical University at Darmstadt, Germany.

150. Sud, A. et al., 2007. Real-time navigation of independent agents using adaptive roadmaps. *Proceedings of the 2007 ACM symposium on Virtual reality software and technology VRST 07*, 1(212), p.99.

151. Tajima, Y. & Nagatani, T., 2001. Scaling behavior of crowd flow outside a hall. *Physica A: Statistical Mechanics and its Applications*, 292(1-4), pp.545–554.

152. Teknomo, K., 2002. *Microscopic pedestrian flow characteristics: Development of an image processing data collection and simulation model*. Tohoku University, Japan.

153. Timmermans, H., Hagen, X. & Borgers, A., 1992. Transportation systems,retail environments and pedestrian trip chaining behaviour: Modelling issues and applications. *Transportation Research Part B*, 26(1), pp.45–59.

154. Tolley, R., 2003. Providing for Pedestrians: Principles and guidelines for improving pedestrian access to destinations and urban spaces. *Department of Infrastructure*.

155. Train, K.E., 2003. Discrete Choice Methods with Simulation. *Identity*, 18(3), pp.273–383.

156. Varas, A. et al., 2007. Cellular automaton model for evacuation process with obstacles. *Physica A: Statistical Mechanics and its Applications*, 382(2), pp.631–642.

157. Venuti, F. & Bruno, L., 2009. Crowd-structure interaction in lively footbridges under synchronous lateral excitation: A literature review. *Physics of Life Reviews*, 6(3), pp.176–206.

158. Villamil, M., Musse, SR & Oliveira, L. de, 2003. A model for generating and animating groups of virtual agents. In *Intelligent Virtual Agents*. Springer Berlin Heidelberg, pp. 164–169.

159. Vovsha, P., 1997. Application of cross-nested logit model to mode choice in Tel Aviv, Israel, metropolitan area. *Transportation Research Record*, (970387), pp.6–15.

160. Vovsha, P. & Bekhor, S., 1998. Link-nested logit model of route choice : Overcoming route overlapping problem. *Transportation research record*, (1645), pp.133–142.

161. Wang, G., Yang, J. & Guo, X., 2008. Research and Application of Cross-Nested Logit Model in Traffic Mode-Split. In *Plan, Build, and Manage Transportation Infrastructure in China*. American Society of Civil Engineers, pp. 203–212.

162. Webb, J.D. & Weber, M.J., 2003. Influence of Sensory Abilities on the Interpersonal Distance of the Elderly. *Environment & Behavior*, 35(5), pp.695–711.

163. Weidmann, U., 1993. *Fussgänger der Fussgänger*. ETH Zuerich.

164. Wen, C.-H. & Koppelman, F.S., 2001. The generalized nested logit model. *Transportation Research Part B: Methodological*, 35(7), pp.627–641.

165. Wolff, M., 1973. Notes On The Behaviour Of Pedestrians. In *Peoples In Places: The Sociology Of The Familiar*. Praeger, pp. 35–48.

166. Yang, L.Z. et al., 2005. Simulation of the kin behavior in building occupant evacuation based on Cellular Automaton. *Building and Environment*, 40(3), pp.411–415.

167. Zhang, J et al., 2012. Ordering in bidirectional pedestrian flows and its influence on the fundamental diagram. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(02), p.P02002.

168. Zhang, J et al., 2011. Transitions in pedestrian fundamental diagrams of straight corridors and T-junctions. *Journal of Statistical Mechanics: Theory and Experiment*, 2011(06), p.P06004.

169. Zhao, D., Yang, L. & Li, Jian, 2008. Occupants ' behavior of going with the crowd based on cellular automata occupant evacuation model. *Science And Technology*, 387, pp.3708–3718.

170. Živanović, S., Pavić, A. & Reynolds, P., 2005. Vibration serviceability of footbridges under human-induced excitation : a literature review. *Journal of Sound and Vibration*, 279(1-2), pp.1–74.

List of Publications

A list of papers that have been accepted or submitted based on this research work is presented below:

1. Rashedi Ashrafi, Z. & Ismail, K., 2013. Exploratory data analysis for calibration of pedestrian models. *Proceedings of CSCE 2013 General Conference - Congrès général 2013 de la SCGC*. DIS-45. May 29- June 1, Montréal, Québec, Canada.

2. Rashedi, Z. & Ismail, K., 2013. Investigation of pedestrian behaviour at bottlenecks under high density situations using discrete choice models. *Proceedings of Transportation Research Board (TRB) 93rd Annual Meeting*. 14-2397. under review.

3. Rashedi, Z. & Ismail, K., 2013. A novel model for pedestrian group behaviour as discrete choices. *Proceedings of Transportation Research Board (TRB) 93rd Annual Meeting.* 14-2309. under review.

4. Rashedi, Z. & Ismail, K., 2013. A discrete choice model for pedestrian group behavior. *Journal of Choice Modelling*. under review.