Simulation and prediction of dense crowd dynamics

Huu-Tu Dang

THÈSE

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par : *l'Université Toulouse Capitole (UT Capitole)*

Présentée et soutenue le *16/01/2025* **par :**

Huu Tu DANG

Simulation and prediction of dense crowd dynamics

JURY

Julien PETTRÉ	Directeur de recherche, Université de Rennes	Rapporteur
Julie DUGDALE	Professeure, Université Grenoble Alpes	Rapportrice
Guy THÉRAULAZ	Directeur de recherche, Université Paul Sabatier	Examinateur
Itsuki NODA	Professeur, Hokkaido University	Examinateur
Benoit GAUDOU	Professeur, Université Toulouse Capitole	Directeur de thèse
Nicolas VERSTAEVEL	Maître de Conférences, Université Toulouse Capitole	Co-directeur de thèse

---

École doctorale et spécialité :

EDMITT : Ecole Doctorale Mathématiques, Informatique, Télécommunications de Toulouse

Unité de Recherche :

IRIT : Institut de Recherche en Informatique de Toulouse (UMR 5505)

Directeur(s)/Encadrant(s) de Thèse :

Benoit GAUDOU et *Nicolas VERSTAEVEL*

Rapporteurs :

Julien PETTRÉ et *Julie DUGDALE*

Huu-Tu Dang

Simulation et prédiction des dynamiques de foules denses

Encadrants **Benoit Gaudou**, Professeur, Université Toulouse Capitole
Nicolas Verstaevel, Maître de Conférences, Université Toulouse Capitole

L'augmentation significative des événements de rassemblement de masse ces dernières années pose le besoin essentiel de simulations et de prédictions réalistes des mouvements de foules denses avant de mettre en œuvre d'autres applications telles que le contrôle de la sécurité ou la gestion des risques. Cependant, différents phénomènes de foule émergent généralement à certains niveaux de densité alors que chaque approche de modélisation est couramment utilisée pour simuler un ou plusieurs phénomènes de foule spécifiques. Cette thèse étudie l'utilisation de facteurs liés à la densité dans des approches hybrides pour améliorer la simulation et la prédiction des mouvements de piétons dans des situations de forte densité. Elle s'articule autour de deux axes principaux : améliorer la simulation des piétons dans divers scénarios de foule et améliorer le réalisme des prédictions (provenant d'algorithmes d'apprentissage profond)à de trajectoire des piétons dans des scénarios de foules denses.

La première amélioration se concentre sur le développement d'un framework à base d'agents qui permet aux agents piétons de changer dynamiquement de modèle de comportement en réponse aux changements de densité locale. La capacité de basculer entre les modèles est illustrée par le couplage hybride d'un modèle mésoscopique et d'un modèle microscopique pour simuler un scénario d'évacuation d'une foule dense et de grande taille lors de la Fête des Lumières à Lyon, en France. Le modèle hybride est ensuite calibré à l'aide d'un algorithme génétique qui utilise des données réelles sur le flux de piétons extraites d'enregistrements vidéo réalisés durant le festival de 2022. De plus, une analyse de sensibilité locale est effectuée pour évaluer l'impact de chaque paramètre sur les résultats des simulations. Les résultats simulés par le modèle hybride montrent qu'il peut capturer qualitativement les tendances générales des données réelles sur le flux de sortie. De plus, une comparaison des performances entre le modèle hybride et différentes combinaisons de modèles est proposée, montrant que le modèle hybride peut simuler efficacement les piétons dans des scénarios de densité variable tout en des meilleures performances en termes de temps de calcul quand le nombre d'agents devient grand.

La deuxième amélioration consiste à proposer un modèle de réseau de neurone (Social-LSTM) étendu en intégrant un terme de gestions des collisions dans la loss fonction d'entraînement pour résoudre le problème de l'apparition de collisions irréalistes dans la prédiction des trajectoires

des piétons. Le terme de gestion des collisions est basé sur l'énergie d'interaction du temps de collision avec les piétons voisins, un concept appliqué avec succès pour modéliser les interactions entre les piétons dans des recherches antérieures. Un facteur de pondération $\lambda \geq 0$ est utilisé pour ajuster l'influence du terme de diminution de collisions. Les expériences menées sur divers jeux de données empiriques indiquent que l'augmentation de λ réduit considérablement les collisions et les chevauchements de piétons dans les trajectoires prédites tout en maintenant une erreur basée sur la distance relativement stable. Ces résultats préliminaires suggèrent que notre modèle représente une approche hybride prometteuse pour la prédiction réaliste des trajectoires des piétons, en particulier dans les scénarios de foule.

Huu-Tu Dang

Simulation and prediction of dense crowd dynamics

Thesis Supervisors **Benoit Gaudou**, Full Professor, Université Toulouse Capitole
Nicolas Verstaevel, Associate Professor, Université Toulouse Capitole

The significant increase in mass-gathering events in recent years poses the essential need for realistic simulation and prediction of dense crowd movements before implementing further applications such as safety control or risk management. However, different crowd phenomena typically emerge at certain density levels while each modeling approach is commonly used to simulate only one or several specific crowd phenomena. This thesis investigates the use of density-related factors in hybrid approaches to improve the simulation and prediction of pedestrian movements in high-density situations. The investigation is divided into two primary focuses: improving the simulation of pedestrians across various crowd scenarios and enhancing the realism of pedestrian trajectory predictions using the deep learning approach in crowded scenarios.

The first improvement focuses on developing a comprehensive agent-based framework that allows pedestrian agents to dynamically switch between models in response to changes in local density. The ability to switch between models is demonstrated through the hybrid coupling of a mesoscopic model and a microscopic model to simulate a large, dense crowd exit scenario in the Festival of Lights in Lyon, France. The hybrid model is then calibrated using a genetic algorithm that utilizes real-world pedestrian outflow data extracted from video recordings of existing crowds at the festival. Furthermore, a local sensitivity analysis is performed to evaluate the impact of each parameter on the simulation outputs. The results simulated by the hybrid model can qualitatively capture the general trends of the actual outflow data. Additionally, a performance comparison between the hybrid model and different model combinations is provided, showing that the hybrid model can effectively simulate pedestrians in varying density scenarios while maintaining computational efficiency.

The second improvement involves proposing an extended Social-LSTM neural network model by integrating a collision loss term into the training loss function to address the issue of unrealistic collision behavior in pedestrian trajectory prediction. The collision loss term is based on time-to-collision interaction energy with neighboring pedestrians, a concept successfully applied to model interactions between pedestrians in previous research. A weight factor $\lambda \geq 0$ is used to adjust the influence of the collision loss term. Experiments on various empirical

pedestrian datasets indicate that adding λ significantly reduces collisions and pedestrian overlaps in predicted trajectories. However, the effectiveness of λ varies across various datasets, showing different trends of evaluation metrics. Therefore, the optimal λ value must be determined for each specific dataset. In general, the best value λ is chosen to balance distance-based accuracy and collision reduction. These preliminary results suggest that our proposed model represents a promising hybrid approach for realistic pedestrian trajectory prediction, particularly in crowded scenarios.

Acknowledgement

Time flies faster than I thought! It feels like just yesterday when I said “challenge accepted” and decided to pursue a PhD, and now my journey is about to end. Throughout this journey, I have received countless support and assistance from many people. I would like to write these grateful words to them; without their help, this thesis would not have been completed.

First and foremost, I would like to express my deepest gratitude to my supervisors, Prof. Benoit Gaudou and Assoc. Prof. Nicolas Verstaevel, who always listened carefully to my ideas and provided invaluable advice for my research. I have learned so much from them, not only from their research but also from their calmness and empathy. I know that I have missed many deadlines, but they were always smiling and giving me constant encouragement throughout my three years of PhD.

I would like to give enormous thanks to the members of the MADRAS project: Prof. Antoine Tordeux, Assoc. Prof. Alexandre Nicolas, Dr. Mohcine Chraibi, Raphael Korbmacher, Oscar Dufour, and Jakob Cordes, with whom I had many interesting discussions that were very helpful for my research and academic development. A special thanks to Prof. Antoine Tordeux and Raphael Korbmacher for the amazing time we collaborated at the University of Wuppertal.

I would like to extend my sincere thanks to Dr. Jullien Pettré and Prof. Julie Dugdale for accepting the invitation to be the reviewers of this thesis, and to Prof. Itsuki Noda and Dr. Guy Théraulaz for being the examiners of the jury. I really appreciate their careful reading and valuable comments, as well as their attendance at the jury.

Last but not least, I would like to thank my family; their love is beyond what can be described in words. Also thanks to my friends, Vu Thi Cam Nhung, Zenith Arnejo, Trinh Hoang Duc, Tran Tuan Duong, Pham Ngoc Quang Anh, and Dao Van Hoang for their priceless mental support. Thank you all!

Contents

General Introduction **1**

Contributions 2

Manuscript Organization 3

1 Context **5**

1.1 Introduction 6

1.2 Background 7

1.2.1 General concepts 7

1.2.2 Phenomena 9

1.3 Research questions 10

1.4 The objectives of this thesis 11

2 Literature review **12**

2.1 Introduction 13

2.2 Related work 14

2.3 Methodology 16

2.3.1 Articles collection 16

2.3.2 Results 17

2.4 Analysis of pedestrian simulation models 20

2.4.1 Strategic-level modeling 22

2.4.2 Tactical-level modeling 22

2.4.3 Operational level modeling 25

2.4.4 Comparison of operational level models 41

2.5 Future directions 52

2.6 Summary and discussion 54

3 Data Collection **57**
3.1 Festival of Lights 58
3.2 Experimental design 60
3.3 Methodology 61
3.3.1 Trajectory extraction method 61
3.3.2 Outflow calculation method 65
3.4 Dataset 66
3.4.1 Trajectory data 66
3.4.2 Outflow data 68
3.5 Validation 69
3.6 Data availability 70
3.7 Conclusion 71

4 HyPedSim – A Hybrid Pedestrian Simulation Framework **73**
4.1 Introduction 74
4.2 HyPedSim Framework 76
4.2.1 General Overview 76
4.2.2 Agent-Based Model for Multi-Level Behaviour 77
4.2.3 Pedestrian Activity Diagram 80
4.3 Application to the Festival of Lights 80
4.3.1 Simulation scenario 82
4.3.2 Pedestrian Simulation Models 83
4.3.3 Simulation Details 88
4.3.4 Model Calibration 88
4.3.5 Calibration Results 92
4.3.6 Model Validation 94
4.3.7 Sensitivity Analysis 94
4.4 Performance Analysis 98
4.5 Conclusions and Discussion 101

5 TTC-SLSTM: Pedestrian Trajectory Prediction Using Time-to-collision Interaction Energy **104**
5.1 Introduction 105
5.2 Problem formulation 106

5.3 Methodology . . . 107
5.3.1 Time-to-collision . . . 107
5.3.2 Interaction energy . . . 110
5.3.3 Social-LSTM . . . 110
5.3.4 TTC-SLSTM . . . 112
5.4 Experiments . . . 114
5.4.1 Datasets . . . 115
5.4.2 Evaluation metrics . . . 116
5.4.3 Implementation details . . . 118
5.4.4 Results . . . 118
5.5 Conclusion . . . 125

6 Conclusions and Discussions **127**
6.1 Conclusions . . . 128
6.2 Scientific contributions . . . 130
6.3 Discussions . . . 131

Glossary **134**

Own Bibliography **135**

Bibliography **137**

List of Figures **161**

List of Tables **163**

General Introduction

The frequency of large-scale events has increased globally over the years, with many reaching high-density levels of crowds. At these levels, the locomotion of most pedestrians in the crowds is significantly restricted due to limited walking space, leading to an increase in complex interactions among pedestrians. For instance, at mass-gathering festivals or crowded peak-hour metro stations, pedestrians tend to move slowly and physically interact with their neighbors. Therefore, the dynamics of pedestrian flow vary dramatically depending on crowd density. This poses an essential task of investigating the role of crowd density in simulating and predicting how pedestrians move and interact with each other in these situations before implementing other measurements, such as safety control or risk assessments. Computer modeling is an effective tool for this task, especially with recent advancements in computational power that allows modellers to simulate and predict crowd dynamics at a large scale. Through the computer modeling approach, researchers can test various scenarios of crowd dynamics under different environmental conditions and model parameters.

Numerous approaches for crowd simulation and prediction exist in the literature, each exhibiting its own advantages and limitations. Specific models are typically recognized for their efficacy in addressing particular characteristics of crowds. Consequently, it is still questionable whether a single model can effectively capture various kinds of crowd dynamics in heterogeneous environments where crowd density changes both temporally and spatially or whether using a hybrid approach may be necessary to have more comprehensive simulations and predictions. This question motivates this thesis to investigate how to incorporate density-related factors into two categories of models, the prediction model and simulation model, to simulate and predict dense crowd dynamics.

Contributions

The novel contributions of this thesis are as follows:

- The first contribution is a comprehensive review of state-of-the-art techniques used in the simulation and prediction of dense crowds. A thorough examination and comparative assessment of modeling techniques for pedestrian behavior across different decision-making levels are conducted using evaluation criteria specifically tailored to high-density scenarios. Through this analysis, the advantages and disadvantages of existing approaches for high-density situations are highlighted together with potential research directions.

- The next contribution is a public release of a high-density pedestrian dataset, with a maximum density of 2.6 $\mathrm{ped}/\mathrm{m}^2$, obtained from mass-gathering events during the Festival of Lights 2022 in Lyon, France. This data collection was conducted in collaboration with members of the Franco-German MADRAS project team. The dataset addresses the current lack of publicly available data in high-density real-world scenarios, as most existing dense crowd data originates from controlled laboratory experiments. Subsequently, this dataset is used to improve two types of models: simulation models and prediction models.

- The third contribution relates to the simulation model, which is a novel agent-based framework for coupling models to simulate dense crowd dynamics. First, we propose a generic architecture to couple models that enables agents to dynamically switch their models based on local crowd density. This architecture is designed to accommodate any model as long as it satisfies specific input-output conditions. Next, the ability to combine different models is demonstrated by an example of coupling microscopic and mesoscopic models for a case study of a mass-gathering event – the Festival of Lights in Lyon, France. We design a genetic algorithm to calibrate and validate the framework using empirical data collected from the festival. Simulation and prediction results are analyzed and evaluated to highlight the advantages and drawbacks of the framework.

- The final contribution involves improving prediction models, which are neural network models for predicting pedestrian trajectories in dense crowds. Neural network models have newly emerged in the last decade, and it is identified in the aforementioned systematic review that these models exhibit unrealistic predictions with many collisions in high-density scenarios. To address this, the Social-LSTM [11] model is extended by integrating a collision loss component into the training function. The collision loss is derived

from interaction energy that quantifies the local density based on time-to-collision with neighboring pedestrians.

Manuscript Organization

The manuscript is organized as follows:

- **Chapter** 1 provides the context for this study. Firstly, it introduces the common concepts and terminologies used in pedestrian simulation and prediction, especially in crowded environments. Next, the chapter addresses the problem state and the motivation of this study, followed by the research question and the objectives of the thesis.

- **Chapter** 2 presents a systematic review that analyzes modeling techniques for simulating and predicting dense crowds. The chapter starts with the methodology used to collect a huge number of articles related to the modeling of multi-level decision-making pedestrian behavior in high-density scenarios. Subsequently, the models developed in these articles are analyzed and evaluated to highlight their advantages and disadvantages, using various evaluation criteria suitable for high-density situations. Finally, the chapter concludes by proposing future research directions in dense crowd simulation and prediction.

- **Chapter** 3 describes the high-density data collection conducted during the Festival of Lights 2022 in Lyon, France. This chapter first outlines the experimental design and the data collection methodology. It then presents data records obtained and analyzes these data, as well as the technical validation of the data.

- **Chapter** 4 proposes an agent-based framework for coupling simulation models. The chapter begins with the design of a generic architecture for the framework aimed at integrating various simulation models. This architecture is designed to allow pedestrian agents to dynamically modify their model based on local crowd density. Next, the ability to couple models of the framework is demonstrated through a real-world case study involving the simulation of large, dense crowds exiting during the Festival of Lights. In the last section, the performance of the framework is evaluated, and the validation and sensitivity analysis are conducted using empirical high-density data.

- **Chapter** 5 proposes the extended Social-LSTM model [11] to predict more accurately pedestrian trajectories in densely populated environments. This extension integrates a collision loss term into the training function. The collision loss component is based on

interaction energy using time-to-collision of neighboring pedestrians. In the experiment, the proposed model is trained and tested on both low-density and high-density datasets. Finally, the prediction results are evaluated using not only distance-based accuracy metrics but also newly proposed collision metrics.

- Finally, conclusions of results and discussion on limitations and perspectives are presented in **Chapter** 6.

Declaration of Generative AI and AI-assisted technologies in the writing process: During the preparation of this manuscript, the author(s) used GPT-4 in order to improve readability and language. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the manuscript.

Chapter 1

Context

Contents

1.1 Introduction 6

1.2 Background 7

1.2.1 General concepts 7

1.2.2 Phenomena 9

1.3 Research questions 10

1.4 The objectives of this thesis 11

This chapter provides the context for the study by first introducing the problem that motivates this research. It then presents the common concepts and terminologies used in crowd simulation and prediction. Finally, it outlines the research questions and objectives of the thesis.

1.1 Introduction

The early 2000s witnessed the emergence of several influential models [12, 13, 14, 15] for simulating pedestrian behavior. Since then, this field has grown significantly over more than two decades with considerable effort dedicated to developing new techniques as well as variations of these models to simulate and predict pedestrian behavior in more complex situations. Pedestrian simulation and prediction in high-density scenarios is one of the typical cases. However, realistic simulation and prediction of large dense crowds (**with a density greater than** $2\ \mathbf{ped/m^2}$) remains a challenging task for several reasons.

The first reason is that pedestrian behavior varies with crowd density [16, 17]. For example, pedestrians can walk freely and sometimes avoid collisions with others in low-density settings. However, in high-density situations, they may have to follow people in front of them to navigate through dense crowds. This difference in pedestrian behavior is because personal space plays a vital role in shaping pedestrian behavior, and it is significantly reduced in crowded environments due to numerous surrounding neighbors. Consequently, pedestrians must adapt their behavior to the new situation. Furthermore, complex interactions, such as physical interactions, are more likely to occur in crowded situations. Therefore, many models exhibit known limitations when applied in different high-density scenarios. For example, the Social Force Model generates abnormal oscillations [18], the Velocity Obstacle Model causes congestion in high-density bidirectional flow [19], the data-driven models yield unrealistic collisions in predictions [7, 9].

In addition, the validation of simulation and prediction models for dense crowds remains an open question. While most models have been calibrated for low-density situations [20], there exists a notable research gap in obtaining empirical data in high-density conditions. The acquisition of empirical data in densely populated real-world events is particularly challenging due to the difficulties in accurately extracting pedestrian data from recordings [21]. On the other hand, laboratory data from controlled experiments has often been found to be expensive and lacks realism in several contexts, such as evacuations, because of ethical constraints and the balance between realism and potential risks to participants [17].

These aforementioned challenges in dense crowd simulation motivate this research, which is a part of MADRAS (Multi-Agent modelling of Dense cRowd dynAmicS: Predict & Understand) [1]. It is a French-German project funded by Agence Nationale de la Recherche (ANR) and

Deutsche Forschungsgemeinschaft (DFG) for a 3-year grant. This project aims to develop accurate and reliable agent-based models to predict and understand dense crowd dynamics (densities from 2 to 8 $\mathrm{ped}/\mathrm{m}^2$). Two modeling approaches are primarily pursued, including neural network models and physics-based models. The novelty of the project lies in the focus on the high-density regimes of crowds, an aspect often underrepresented in current research literature.

Partners from four universities with different backgrounds participate in the project, with two from the French side, Toulouse Capitole University, Claude Bernard University Lyon 1, and the other two from the German side, University of Wuppertal and University of Cologne. Each PhD candidate has a three-month cross-institute research visit to collaborate with teams from the other side. The author spends this period at the University of Wuppertal working with Antoine Tordeux and Raphael Korbmacher. The collaboration focuses on developing a deep learning approach to predict pedestrian trajectories in crowded situations.

1.2 Background

Crowd simulation and prediction is an interdisciplinary research field that aims to understand, simulate, and predict pedestrian behaviors in various environments. This field has been applied in numerous domains, such as urban planning [22], computer graphics [22, 23], evacuation [24], safety science [25], etc. This section presents the primary concepts and terminologies used in crowd simulation. A glossary of similar terminologies can also be found in [20, 26, 27].

1.2.1 General concepts

From the perspective of crowd simulation, a **pedestrian** is defined as a person moving on foot in a publicly accessible area [27]. The **velocity** of a pedestrian refers to its walking speed in the direction of movement. When each pedestrian moves, they create a **trajectory** which is a sequence of pedestrian positions over time [28].

Pedestrian behaviors can be classified into different decisional levels based on their objectives and environment. This thesis relies on the classification criteria proposed by Hoogendoorn and Bovy [29] to define **multi-level behaviors** of pedestrians:

- **Strategic level**: at this level, pedestrians formulate a list of desired activities (or destinations, in other words) and establish a timeline for each task that they intend to accomplish during their journey.

- **Tactical level**: pedestrians compute a route to reach their predetermined targets based on their knowledge of the environment. This route can be global or local, depending on their understanding and familiarity with the environment.
- **Operational level**: pedestrians actualize decisions made at the strategic and tactical levels through physical movement while managing local interactions, such as navigating among neighbors or avoiding collision with obstacles.

The classification of these three hierarchical levels provides a holistic perspective of pedestrian behaviors by considering both long-term and short-term decision-making behaviors rather than solely focusing on local movements of pedestrians.

Situations containing a large number of pedestrians can form a crowd. There are many definitions of a crowd in the literature. In this study, we use the definition provided in the review conducted by Duives et al. [20]:

> "A **crowd** is a large group of individuals ($N \geq 100P$) within the same space at the same time whose movements are for a prolonged period of time ($t \geq 60s$) dependent on predominantly local interactions ($k \geq 1\ \mathrm{ped/m^2}$)."

The most fundamental aspects of a crowd are often captured by the **fundamental diagram**, which plots the functional relationships between three quantities: density, speed, and flow.

The **density** of a crowd is calculated as the average number of pedestrians per space unit at a given time. Although the definition of high density may vary across domains and applications, this thesis defines the high-density level using the condition of pedestrian movement in the crowd. Accordingly, the **density level is considered high if it is greater than** $2.0\ \mathbf{ped/m^2}$. This threshold is chosen because densities at or above this level typically restrict all pedestrian movements [30]. The density of crowds in some real-world case studies [31] can reach extremely high values up to $9\ \mathrm{ped/m^2}$.

On the other hand, the **flow** of a crowd is the rate of pedestrians passing a specific location or area per time unit. Previous studies have shown that crowd flow starts to decrease at densities exceeding $2.8\ \mathrm{ped/m^2}$ [32].

A common way to model pedestrians' behaviors is using multi-agent systems. From this perspective, pedestrians are considered as agents that are simulated in computer systems. There are also various definitions of an agent, this study uses the one that is commonly accepted:

> "An **agent** is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its design objectives" [33].

An agent can interact with other agents as well as with the environment. The **environment** represents the spatial settings in which agents are situated and perform actions at each simulation step. It is characterized by various elements such as obstacles, buildings, walkable spaces, facilities, etc. Depending on the experiments, the environment can vary from simple corridors to complex shapes represented by shapefiles.

1.2.2 Phenomena

Several emergent pedestrian behaviors and collective phenomena have been observed in high-density situations. The following describes the main pedestrian behaviors and crowd phenomena used in this study.

Group behavior. When walking in social groups, such as friends or family, pedestrians exhibit group behavior to stay close to one another. This socially coordinated behavior is influenced by the movements of other group members and aims to maintain a certain level of cohesion within the group. The group's structure depends on the number of members and the density of the surrounding area. In low-density environments, people in groups typically walk side by side, but in crowded areas, they tend to form a V-like shape [34]. Modeling group behavior of pedestrians in real-world scenarios is important as many pedestrians in crowds are part of social groups [34, 35].

Lane formation. Lane formation is a phenomenon in which groups of pedestrians walking in the same direction self-organize into lanes. This self-organization minimizes conflicts and interactions with other pedestrians moving in opposing streams, especially in crowded areas. The lane formation phenomenon is typically observed in the continuous bidirectional or crossing flow of pedestrians, such as in long, narrow corridors.

Following behavior. The following behavior, also called leader-follower behavior, occurs when pedestrians follow predecessors who are moving in the same direction. This behavior is likely to emerge in situations where personal space is limited and the route ahead is not clearly visible. As a result, pedestrians tend to follow the movement of leaders to reduce potential interactions and to navigate through dense crowds [36]. From a macroscopic perspective, the localized following behavior of individuals leads to the formation of lanes in crowds.

Physical interaction. Physical interactions usually occur in crowded scenarios where pedestrians lack sufficient space for free movement and tend to interact physically with their neighbors. These interactions are, for example, physical contact, collisions, or pushing. The impact of physical interactions on pedestrian behavior and the dynamics of crowds becomes particularly significant in extremely high-density scenarios, as they are the primary reason for

crowd turbulence [31].

Replan behavior. Replan behavior refers to how pedestrians modify their strategic or tactical plans to adapt to the dynamic state of their surroundings. This behavior is predominantly psychological and reflects the decision-making processes of individuals prior to translating into physical actions. For instance, pedestrians may initially possess a predetermined route to their destination but may choose to replan their planned route or even alter their destination to move away from congested areas.

Shockwaves. Shockwaves in crowd dynamics are wave-like propagation caused by sudden and discontinuous changes in the flow, density, or speed of pedestrians. Shockwaves typically occur in unidirectional crowds at densities above $4.3\ \mathrm{ped/m^2}$ [37] and usually are observed before crowd turbulence happens [31].

Crowd turbulence. At critical density, the crowd becomes extremely compressed with pedestrians experiencing pushing forces from multiple directions [16, 37]. This leads to panic and turbulent behavior of pedestrians, which are the primary causes of falls, stampedes, and related incidents. Crowd turbulence is typically observed when densities exceeding $8.5\ \mathrm{ped/m^2}$ [37].

In summary, different crowd behaviors and phenomena have been observed at certain high-density levels. Crowd flow starts to decrease [32] when density exceeds $2.8\ \mathrm{ped/m^2}$. Stop-and-go waves usually occur [37] at densities higher than $4.3\ \mathrm{ped/m^2}$. Crowd turbulence arises when density is above $8.5\ \mathrm{ped/m^2}$. In some exceptional cases, densities can reach up to $9\ \mathrm{ped/m^2}$ [31]. Therefore, modeling high-density crowds must take into account emergent phenomena corresponding to each density interval to ensure the realism of simulations.

1.3 Research questions

Given the crucial role that crowd density plays in modeling dense crowds, this thesis aims to investigate the use of density-related factors to improve two types of models: prediction models and simulation models. Specifically, this thesis seeks answers to the following research questions related to hybrid approaches taking density factor into account to simulate and predict dense crowd dynamics:

- How to integrate density-related factors into prediction and simulation models to address high-density scenarios?
- How to couple simulation models in a single framework to have more comprehensive simulations of crowd dynamics across various densities?

- How can we combine or switch simulation models in a generic manner? Is a general architecture needed for this purpose? Furthermore, can this general architecture also be adapted to simulate a variety of dynamics, not limited to crowd dynamics?
- For prediction models, is it necessary to develop a new neural network architecture that specifically incorporates density factors to predict pedestrian trajectories in high-density situations or can density factors be integrated into existing neural network models?
- Given density factor integrated into prediction and simulation models, how to calibrate and validate these models based on empirical data in high-density situations?

1.4 The objectives of this thesis

This thesis aims to pursue these key objectives:

- O1: The first objective is to review and understand the strengths and limitations of existing prediction models and simulation models for dense crowds. From that, we identify the current research gaps in the literature and propose directions to improve these models.
- O2: The next objective is to collect high-density data from real-world case studies and use it to calibrate and validate these above models.
- O3: The third objective focuses on simulation models. We aim to design an agent-based framework for coupling different simulation models using the crowd density factor to simulate various crowd phenomena. This framework must be generic enough to allow numerous models to be integrated. Furthermore, the architecture of the framework should be extensible to other dynamics like traffic simulation.
- O4: The final objective aims to improve prediction models by integrating density factors into neural network models to tackle the problem of unrealistic prediction generated in crowded scenarios. Concurrently, the design of novel evaluation metrics is also targeted with the aim of providing more accurate quantification of predictions in high-density conditions rather than only using distance-based metrics.

Chapter 2

Literature review

Contents

2.1 Introduction 13

2.2 Related work 14

2.3 Methodology 16

2.3.1 Articles collection 16

2.3.2 Results 17

2.4 Analysis of pedestrian simulation models 20

2.4.1 Strategic-level modeling 22

2.4.2 Tactical-level modeling 22

2.4.3 Operational level modeling 25

2.4.4 Comparison of operational level models 41

2.5 Future directions 52

2.6 Summary and discussion 54

The previous chapter outlined the context, motivation, and objective of this study, along with related concepts, terminologies, and phenomena commonly used in pedestrian modeling and simulation. It highlighted our goal to investigate the role of density-related factors in developing models for simulating and predicting pedestrian dynamics in crowded situations. Before presenting my own contributions in the following chapters, this chapter presents a systematic review of state-of-the-art techniques for modeling pedestrian behaviors with a primary focus on high-density scenarios. This work has been published in [6] where the author was the primary contributor.

2.1 Introduction

Pedestrian modeling and simulation have been an active research area for several decades, particularly since the early 2000s with the advent of various well-known models ranging from microscopic models [12, 13, 38] such as Social Force Model to macroscopic models [14, 15]. These models can be used to analyze the capacity of infrastructure, for example, in Hajj religious events.

Among all the works about pedestrian modeling and simulation, interest for high-density situations is still new and these situations are less explored. As a consequence, there is only a small ratio among existing reviews on approaches for modeling pedestrian behaviors that focus on this issue. A detailed discussion of these reviews is presented in the next section. In contrast, our systematic review offers the following novel contributions:

- A thorough investigation of various approaches for modeling pedestrian behaviors at different decision-making levels, with investigation starting from the classical model of each approach to its variations to simulate pedestrian behavior in high-density scenarios.
- A comprehensive evaluation and comparison of these approaches using various criteria appropriate for high-density conditions, such as simulated density level and model's ability to simulate pedestrian emergent behavior in crowded situations, among other factors.
- An assessment of the strengths and limitations of these approaches in handling dense crowd characteristics.
- A proposal of potential research directions for future development of pedestrian modeling in high-density conditions.

Given the advent of many well-known models in the last two decades, our review focuses on studies published from January 1, 2000, to December 31, 2023.

This chapter is organized as follows. Section 2.2 provides an overview of existing reviews on pedestrian modeling and discusses the limitations of these reviews in summarizing modeling approaches for high-density scenarios. Section 2.3 details the methodology for selecting the corpus of relevant articles and provides a first quantitative analysis of these articles to highlight emerging trends. Next, Section 2.4 examines, evaluates, and compares the advantages and disadvantages of state-of-the-art approaches for modeling pedestrian behaviors across different decision-making levels in high-density environments using a selected set of criteria. Section 2.5 proposes potential research directions in the future development of dense crowd modeling and simulation. Lastly, the chapter concludes with a summary and discussion in Section 2.6.

2.2 Related work

Numerous surveys and reviews have been conducted to summarize existing modeling approaches for pedestrian simulation. For instance, Helbing and Johansson [26] provided a short overview of the development history of the field of pedestrian modeling and simulation. Furthermore, this study explored the challenges of modeling the self-organized and panic behavior of pedestrians during evacuations and highlighted the effectiveness of social force-based models in addressing these phenomena. Zhou et al. [39] proposed qualitative criteria to evaluate pedestrian modeling approaches based on model capability and performance. Yang et al. [40] carried out a survey to review methods for modeling pedestrian behaviors developed over the last decade. Musse et al. [41] provided an overview of the evolution of pedestrian modeling techniques through different time periods through an extensive collection of articles over the last three decades. Likewise, van Toll and Pettre [42] reviewed the main techniques that have emerged or been improved to simulate pedestrian behavior in the past decade, with the primary focus on microscopic models. Bellomo and Christian [43] provided a comprehensive examination of mathematical models for simulating macroscopic characteristics of crowd dynamics.

Besides these general reviews, some other papers attempt to specific characteristics or phenomena modeled in crowd simulations, such as group dynamics, and hybrid modeling... Cheng et al. [35] carried out a review to examine microscopic techniques capable of modeling pedestrian group dynamics. Lately, Alexandre and Fadratul [44] analyzed the effect of social groups on the overall crowd movement and summarized different approaches to model these social groups. Ijaz et al. [45] conducted a review to investigate hybrid approaches that combine different pedestrian modeling paradigms. Papadimitriou et al. [46] examined models for route choice and crossing behavior in urban settings, highlighting the lack of models capable of simulating complete pedestrian trips. A thorough review was performed by Basu et al. [47]

to investigate various factors affecting the route choice behavior of pedestrians. These factors were then classified into three classes, including pedestrian socio-demographic characteristics, environmental characteristics, and trip characteristics.

In parallel to these surveys mainly focused on models, another set of papers discussed empirical studies and the use of data in pedestrian simulations. Surveys reviewing experimental pedestrian data and empirical approaches can be found in [17, 48]. The studies noted that most empirical datasets were collected under normal conditions, with only a few datasets obtained in emergency and panic scenarios. Almoaid et al. [49] reviewed modeling approaches that have been specifically applied to simulate the Hajj-Tawaf pilgrimage - the world's largest mass gathering. Their findings indicated that microscopic models were commonly used to simulate dense crowds in this case study; however, the highest number of pilgrims simulated in these studies was significantly lower than the actual number of attendees at the event. Feng et al [21] conducted a thorough examination of approaches for collecting pedestrian data. Zhong et al. [50] reviewed various methods for calibrating and validating pedestrian simulation models using empirical data. Korbmacher and Tordeux [28] conducted a comprehensive comparison of the accuracy of neural network algorithms and classical pedestrian simulation models in predicting human trajectories. Their result indicated that neural network algorithms performed more accurate predictions compared to classical pedestrian simulation models.

While there have been many reviews and surveys on pedestrian modeling and simulation, **the majority have focused primarily on regular circumstances rather than high-density situations**. Recent systematic reviews, particularly since 2021, have notably lacked focus on modeling methods that simulate pedestrian behavior in high-density scenarios. More specifically, Musse et al. [41] provided a historical overview of emerging techniques from each time period but did not analyze in detail their technical aspects. van Toll et al. [42] presented a thorough and insightful analysis of prevalent microscopic models for simulating pedestrian behavior; however, their survey lacks further quantitative metrics to evaluate and compare model performance. Korbmacher and Tordeux [28] primarily focused on comparing classical pedestrian simulation models and neural network models in terms of application domains, technical aspects, and distance-based prediction accuracy between in low-density situations. Zhong et al. [50] focused only on examining calibration and validation methods for pedestrian simulation models. Basu et al. [47] mainly investigated approaches for modeling the route choice behavior of pedestrians while not considering local movement and interactions. While Nicolas and Hassan [44] explored methods for modeling social groups across various densities, social groups represent only one facet in the broader spectrum of modeling crowd dynamics in high-density scenarios. The comprehensive review carried out by Dorine et al. [20] in 2013 remains the only study

investigating pedestrian modeling techniques at both low-density and high-density conditions; however, there is a pressing need for covering more recent studies addressing high-density crowds. To address this gap, we conducted a systematic review with a main focus on examining modeling techniques for simulating pedestrian behavior in high-density scenarios.

2.3 Methodology

2.3.1 Articles collection

A large number of relevant articles were collected from two databases: Web of Science and Science Direct. Suitable keywords have been prepared, which address the primary focus on simulating pedestrian behavior in high-density scenarios, to query the databases:

"TS = ((crowd OR pedestrian) simulation model AND (dense OR high-density))"

Here, the phrase "crowd OR pedestrian" ensures the inclusion of all research related to pedestrian crowds, whereas the phrase "simulation model" focuses on the development of models to simulate pedestrian behaviors. Next, the phrase "dense OR high density" highlights the specific interest in high-density contexts, which is the focus of this study. Finally, the "AND" operator connects the key search terms to retrieve relevant articles.

Figure 2.1 presents the flow diagram for collecting articles relevant to the modeling of pedestrian behavior in high-density situations. A systematic search using the aforementioned search keywords was conducted in January 2023 in the two databases. The search was limited to titles and abstracts of journal articles and conference proceedings published between 2000 and 2022. The search resulted in 458 papers in Web of Science and 445 papers in Science Direct, respectively. Furthermore, an additional 152 relevant articles were collected through a thorough check of the reference lists of the initially collected articles. This process involved manually examining the citations within these articles to find studies that align with our focus on high-density pedestrian behavior. Additionally, recent studies published in 2023 were included. Following the exclusion of 249 duplicates, the screening process was applied to a total of 806 articles.

In the screening process, papers are selected when they satisfy the following inclusion criteria:

- The article must present new techniques for modeling pedestrian behaviors in at least one of three decisional levels of pedestrian behavior [29], with a specific focus on high-density contexts.

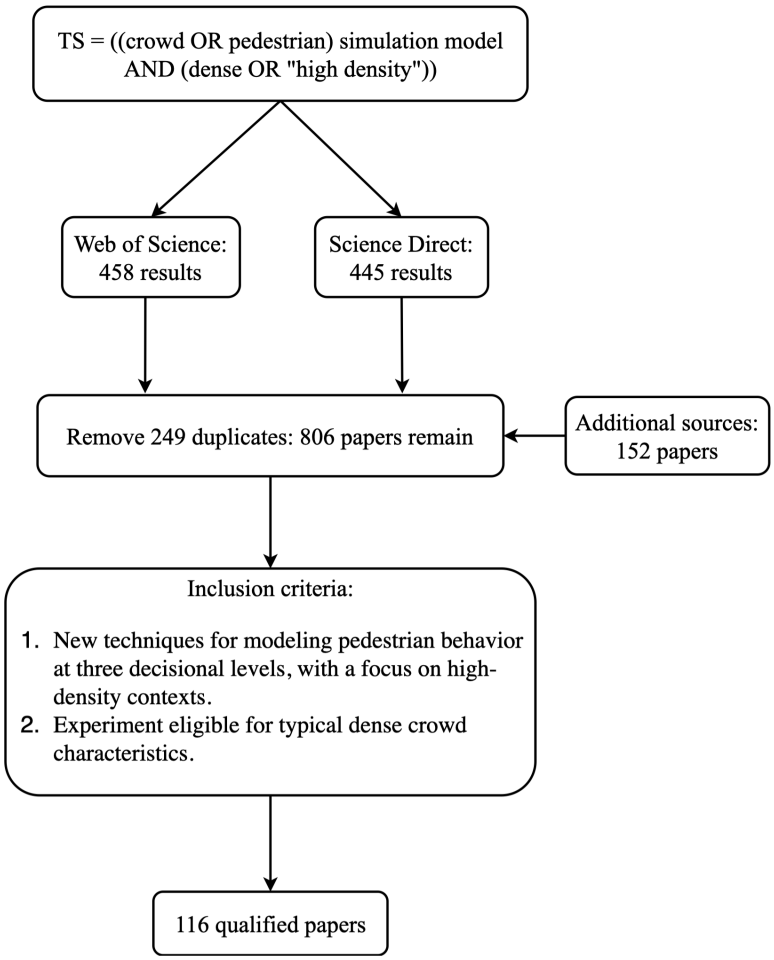


Figure 2.1: Flow diagram for collecting articles relevant to the modeling of pedestrian behavior in high-density situations.

- The simulated situations must contain and/or exhibit the typical high-density crowd characteristics like high-density levels and emergent behavior in crowded situations.

A total of 116 articles finally matched the inclusion criteria, and are analyzed in the next sections. Table 2.3 compares all these models based on the chosen evaluation criteria.

2.3.2 Results

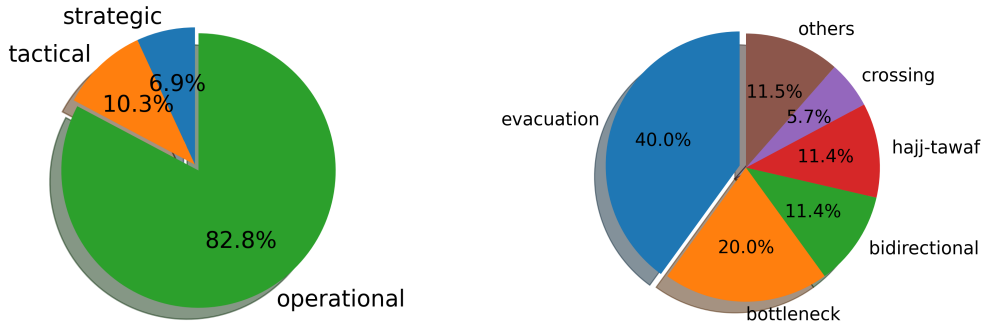
After having collected the corpus of relevant articles, each one is classified into one of the three modeling levels of pedestrian behavior [29]. Note that while one article can exhibit more than one level, it is classified based on its primary focus or main contribution to one

specific level. The main classually es of models within each level of behavior are described in Table 2.1. Tactical-level models include graph-based [51, 52, 53, 54, 55], navigation mesh [56] and potential field [57, 58, 59, 60], while operational-level models consist of Social Force model (SFM) [13, 61], Cellular Automaton (CA) model [12, 38], Velocity Obstacle (VO) model [62, 63], Agent-based model (ABM) [64, 65], data-driven model [11, 66], macroscopic model [14, 15, 67], and hybrid model [68, 69, 70].

Table 2.1: Main classes of models at different levels of pedestrian behavior.

Levels of behavior	*Modeling techniques*
Strategic	–
Tactical	Graph-based [51, 52, 53, 54, 55]
	Navigation mesh [56]
	Potential field [57, 58, 59, 60]
Operational	Forced-based models [13, 61]
	Velocity Obstacle models [62, 63]
	Cellular Automata models [12, 38]
	Agent-based models [64, 65]
	Data-driven models [11, 66]
	Macroscopic models [14, 15, 67]
	Hybrid models [68, 69, 70]

Figure 2.2 presents the percentage of articles contributing to each of the three decisional levels and experiment types, respectively. It is worth noting that although a study may show simulations of pedestrian behavior at different levels, its main contribution is often at one level. The results indicate that most articles develop new techniques that aim to model operational-level behavior, while significantly fewer studies focus on strategic-level modeling. This focus on the operational level is understandable since modeling collision avoidance, the most fundamental challenge in modeling pedestrian behavior, is at the operational level. In addition, experiments in these studies are categorized into several classes, including evacuation [27], bottleneck [20, 27], bidirectional [20], Hajj-Tawaf [25], and crossing [20]. As illustrated in Figure 2.2b, simulations of pedestrian dynamics during emergency evacuations and at bottlenecks are the most popular experiments. Finally, among real-world applications of extremely high density, the Hajj-Tawaf religious event is the main case study.



(a) Distribution of articles contributing to modeling each of the three decisional levels [29].

(b) Distribution of articles according to experiment types.

Figure 2.2: Analysis results on the collected articles.

Figure 2.3 presents the number of yearly publications for dense crowd simulations from 2000 to 2023. The result indicates a notable increasing trend over more than two decades, particularly in the second half of the period. A key factor driving this growth is the recent emergence (around year 2016) of various prominent neural network models for predicting pedestrian trajectories that demonstrate more accurate results compared to classical pedestrian simulation models [71].

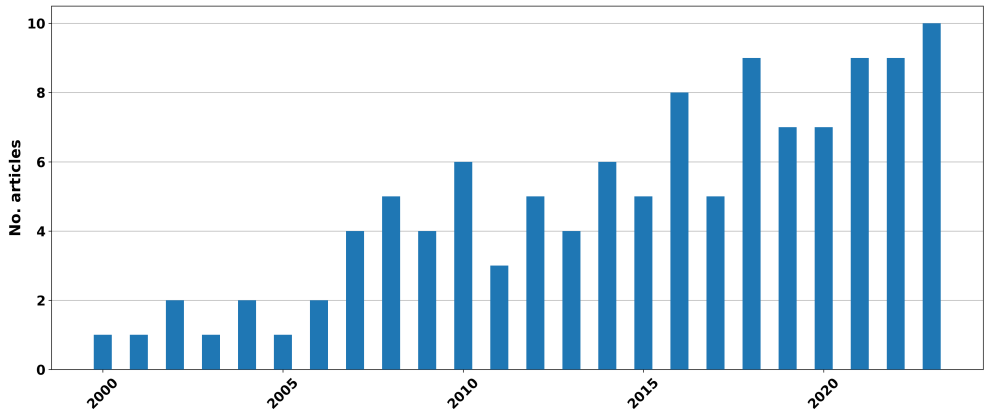


Figure 2.3: Number of publications per year for dense crowd simulations.

Figure 2.4 presents the word cloud extracted from the abstract and authors' keywords of collected articles. The words "pedestrian", "crowd OR pedestrian", and "dynamic" are commonly used in conjunction with "high density" and "evacuation". This suggests that the collection of relevant articles has effectively focused on dense crowd simulation.

Finally, an analysis of the usage of modeling techniques over time is conducted for two

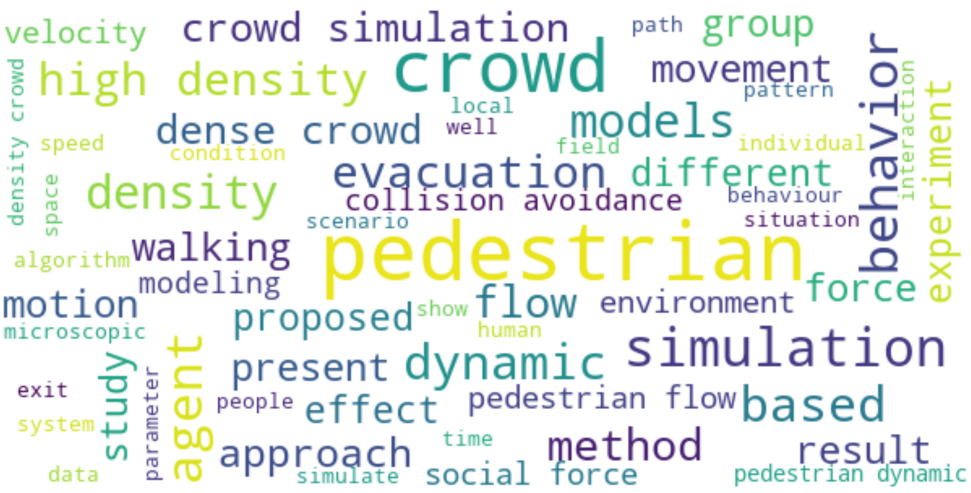
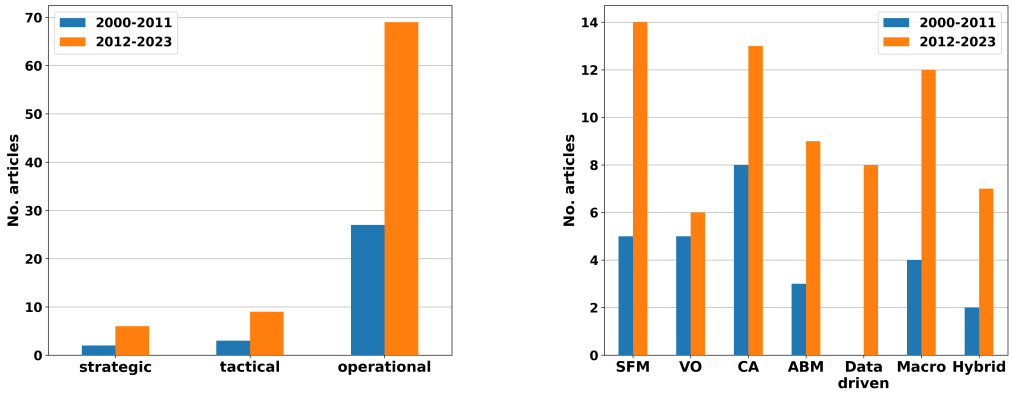


Figure 2.4: Word cloud extracted from the abstract and authors' keywords of collected articles.

distinct periods: 2000 to 2011 and 2012 to 2023. Figure 2.5 presents the variations in the number of studies in the two periods across the three decisional levels and the various modeling approaches at the operational level. The result indicates a general increasing trend at all three levels in the second half period. Nevertheless, a closer look at operational-level models reveals that VO models exhibit only a minor increase in the number of studies in the second half period. On the other hand, the other operational-level models, including SFM, CA model, ABM, macroscopic model, and hybrid model, show a consistent increasing trend. Remarkably, the advent of neural network data-driven models for predicting pedestrian trajectory is relatively new, as supported by recent advancements in computational capacity and available datasets of pedestrian trajectories.

2.4 Analysis of pedestrian simulation models

Modeling pedestrian behaviors in complex environments at three levels, including strategic, tactical, and operational, can comprehensively capture different aspects of pedestrian decision-making and actions [29]. Figure 2.6 presents the classical Perception-Decision-Action schema across these levels. At the strategic level, people determine their desired destination, such as choosing a specific store to visit in a shopping mall. The destination determined at this level serves as the input for the tactical level, where pedestrians incorporate their environmental knowledge with their chosen target to formulate a global or local route that guides them toward their destination. Path planning algorithms are typically used to compute the best route in this



(a) Three decisional levels. **(b)** Operational-level models.

Figure 2.5: Variations in the number of studies over the two periods: 2000 – 2011 and 2012 – 2023.

step. Next, the selected route is implemented at the operational level as a series of waypoints or preferred velocities for pedestrian simulation models to simulate the local movement and interactions of pedestrians. Furthermore, pedestrians may also adjust their initial plans, including their destination or tactical route, in response to dynamic changes in the surrounding environment. In the following subsections, strategic, tactical, and operational-level models are thoroughly examined and compared to identify their strengths and limitations in simulating pedestrian behavior in high-density conditions.

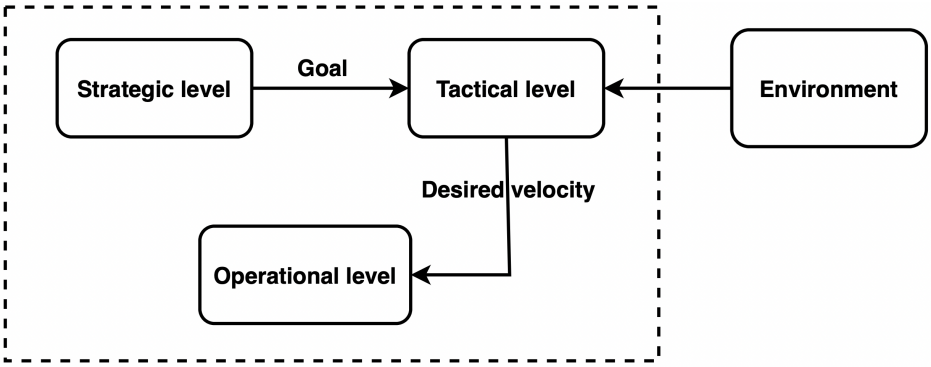


Figure 2.6: The classical Perception-Decision-Action schema.

2.4.1 Strategic-level modeling

Strategic level modeling, often referred to as *goal selection* [72], is a crucial aspect of pedestrian behavior modeling that focuses on how pedestrians select their destinations or targets. This selection process is primarily driven by psychological interests and individual motivations. Although pedestrian modeling has been broadly studied and applied in numerous applications, modeling pedestrian behaviors at the strategic level has received relatively less attention in the literature compared to strategic and tactical levels, as presented in Figure 2.2. This limited focus on strategic-level modeling can be attributed to several reasons. Primarily, researchers have concentrated on modeling pedestrian behaviors at the tactical and operational levels, which include most pedestrian activities. Next, the shortage of empirical datasets for strategic-level decision-making poses a significant challenge to calibrate and validate strategic-level models [73]. Finally, strategic-level models are more related to modeling human interests, which requires insights from psychology, and has often been investigated separately from the field of pedestrian modeling.

Strategic-level models simulate the goal selection behavior of pedestrians, which generates a target for each pedestrian. These models are predominantly investigated in emergency scenarios, where pedestrians typically identify an exit to escape quickly and safely. Among these models are logit-based discrete choice models [74, 75], genetic-algorithm models [76], exit choice models [73, 77], interest function models [78], and game theory-based models [79], leader-follower exit-choice models [80].

2.4.2 Tactical-level modeling

Tactical-level models, also known as *path planning* [72], simulate pedestrian behavior to prepare navigation strategies to move from current positions to intended destinations using environmental insights. While simulations in simple scenarios like long corridors or bottlenecks can be sufficiently represented using only operational-level models, tactical-level modeling is essential in large, complex environments with numerous obstacles. Without tactical considerations, simulated pedestrians may get trapped in local areas within such environments. Furthermore, a fundamental assumption is that pedestrians own a cognitive map of their environment [81], which they can use to plan routes to their destinations.

Path planning algorithms have been effectively used in robotic navigation [82] as well as agent navigation in virtual environments. It can primarily be classified into two main approaches: global path planning and local path planning:

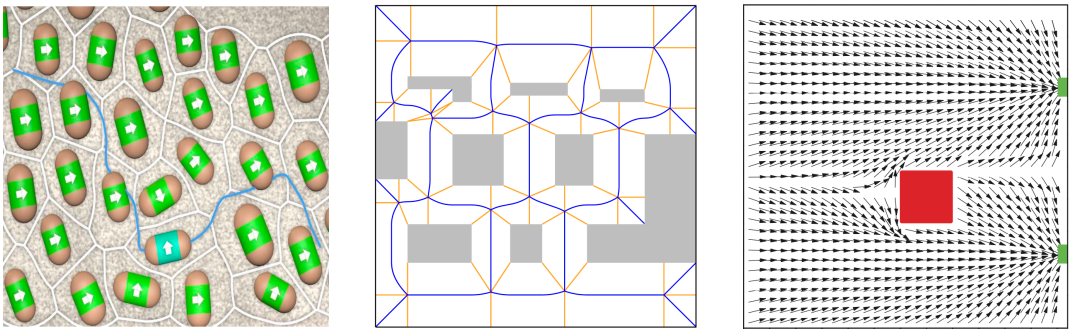
- Global path planning algorithms use the entire environmental map to compute global

paths from the origin to the target. Graph-based models [83, 84] and navigation mesh models [85] are the two main models of this approach. Graph-based models [83, 84] abstract the environment into a graph network of interconnected nodes and edges, where nodes signify critical locations or points, and edges denote the feasible routes between these nodes. Conversely, navigation mesh models [85] partition the environment into polygonal zones representing walkable areas. In general, path search algorithms are typically applied in these models to compute the shortest or fastest path for each individual.

- Local path planning algorithms like potential fields [86] compute local routes based on perceptual information from pedestrians' surrounding environment without requiring a full environmental map.

While tactical-level models traditionally rely on using static environmental data, such as obstacles and walkable areas, to compute guiding paths in low-density scenarios, this approach works well in simple environments but is inadequate in densely populated environments where the movement of other agents significantly impacts navigation. This limitation of static path planning in crowded scenarios can be exemplified by the "faster-is-slower" effect [26, 87], where a large number of pedestrians choosing the same shortest path based on static information, leading to increased travel time and potential congestion. To tackle this issue, dynamic crowd-related information such as crowd density or estimated space availability must be taken into account dynamically for more effective and adaptable path planning[29]. Similar techniques [88, 89, 90] are being applied to robotic navigation in high-density scenarios.

Global path planning commonly incorporates dynamic crowd-related data to dynamically adjust graph weights. This allows path search algorithms to return adaptive routes that reflect current crowd state information. Various studies have been proposed to improve global path planning with dynamic information updates. For instance, Stubenschrott et al. [54] have introduced local crowd density to estimate perceived travel time along graph edges, whereas Sud et al. [53] have incorporated crowd density factors for continuous updates to roadmap weights. The Principle of Least Effort, proposed by Guy et al. [91], computes the best energy-efficient paths for pedestrian navigation by updating graph weights according to average pedestrian flow speeds. Stüvel et al. [92] have integrated a clearance factor to Voronoi edges (as presented in Figure 2.7a) to enable pedestrians to choose the best path for navigating through densely populated environments. To model crowd movement in emergency scenarios, Liu et al. [55] have proposed incorporating emotional factors and perilous field concepts into Delaunay triangulation representations. Van Toll et al. [56] have proposed partitioning the environment into navigation meshes using the medial axis and Explicit Corridor Map (as illustrated in Figure 2.7b), in which



(a) Graph network in [92] **(b)** Navigation mesh in [56] **(c)** Navigation field in [59]

Figure 2.7: Different tactical-level models for pedestrian navigation: (a) Graph network generated based on Voronoi edges (white lines); (b) Navigation mesh constructed using the medial axis (blue lines) and orange lines connecting the medial axis vertices to proximal obstacles; (c) Navigation field describing preferred moving directions of crowds across different local areas.

crowd density in each mesh is updated when pedestrians move between meshes. However, in large-scale environments like cities, using the global path-planning approach can lead to computational challenges, especially during path searching. To address this issue, hierarchical graph structures [52, 51] have been proposed to decrease the computation costs of path planning. Given dynamic information integrated into graph networks, Stubenschrott et al. [54] have enabled pedestrians to search for a better path as soon as they reach a node with multiple path options. In contrast, Van et al. [56] have implemented regular updates of paths for pedestrians after a specific time interval.

In contrast, local path planning considers the assumption that pedestrians have limited visibility and memory constraints, leading them to make tactical decisions based on their intermediate surroundings rather than having full knowledge of the entire environment. Zhang et al. [57] have introduced a perceived potential field, which is extended from earlier potential field [86], to simulate pedestrian tactical behavior with restricted visibility. Jiang et al. [59] have developed a local dynamic navigation field (as illustrated in Figure 2.7c) by solving an Eikonal-type equation on rectangular cells. Sun and Liu [60] have suggested a model that combines a local density navigation field with an equipotential field to simulate pedestrian evacuation in emergency scenarios. Hoogendoorn et al. [58] have incorporated density-gradient dependent terms into the local path selection behavior of pedestrians, which leads to the pedestrian tendency to avoid congested regions.

Table 2.2 shows a comparison of the characteristics of tactical-level models for simulating

pedestrian navigation behavior in densely populated environments. In general, the effectiveness of these models is influenced by various aspects, including environmental complexity, the number of agents, and the algorithms themselves. Global path planning models are particularly beneficial for computing complete routes in large-scale environments where the full environmental map is known. These models are typically applied to initialize paths for a large number of agents using static environmental information. Nevertheless, their computational demands can increase substantially when simulating the replan behavior of pedestrians, particularly for large crowds.

Conversely, local path planning models are typically employed in situations where pedestrians do not perceive full information about the environment or when the crowd dynamics are highly volatile. The benefit of these models is their capacity to calculate short-term, adaptable routes that reflect local changes in crowd characteristics. However, they incur greater computational costs due to frequent updates needed to address the latest changes in crowd characteristics. In conclusion, selecting between global and local path-planning models for simulating pedestrian tactical behavior in high-density situations requires careful consideration of various factors such as environmental complexity, computational resource availability, and the desired level of realism in simulations of pedestrian tactical behavior.

2.4.3 Operational level modeling

Pedestrian simulation models at the operational level aim to simulate the local movement of pedestrians by calculating their velocities for the next simulation steps. These models can communicate with strategic- and tactical-level models through preferred targets and velocities, respectively. Various operational-level models have been proposed to simulate pedestrian behavior across diverse scenarios. They can be broadly classified into three main categories based on the modeling scale:

- **Microscopic models**: these models focus on simulating the local behavior of individual pedestrians and individual-level interactions among pedestrians. Microscopic models are further categorized into **theory-driven models** and **data-driven models**. Theory-driven models, including **Social Force Models (SFMs)**, **Velocity Obstacle (VO) models**, **Cellular Automata (CA) models**, and **Agent-Based Models (ABMs)**, are grounded on fundamental principles and hypotheses. They are typically built for a large range of case studies and aim to be as generic as possible. The application to a specific situation requires the calibration of parameters for realistic simulations. Conversely, data-driven models learn pedestrian behaviors implicitly from empirical pedestrian datasets through neural network architectures. As a consequence, they are built for a single application.

Table 2.2: Comparison of the characteristics of tactical level models for simulating pedestrian tactical behavior in high-density scenarios (CT: Construction time, MUT: Map update time, PST: Path search time, NA: Number of agents, RT: Run time).

Articles	Class	Environment	CT (ms)	MUT (ms)	PST (ms)	NA	RT (fps)
Lamarche et al. (2004) [51]	Graph	City: $1.3km \times 1.3km$, 2600 obstacles	–	–	25(2000)[1]	2000	10
Paris et al. (2006) [52]	Graph	Lazare train station, France	–	11	–	2000	7
Sud et al. (2008) [53]	Graph	Tradeshow: 511 booths	–	5.5	7(1000)	1000	22
Stubenschrott et al. (2014) [54]	Graph	U2 subway station, Austria	–	–	–	–	–
Stuvel et al. (2017) [92]	Graph	Small room	–	–	–	20-30	–
Liu et al. (2018) [55]	Graph	Big maze	150	–	–	2000	$\approx$ 17
Van Toll et al. (2012) [56]	Nav mesh	City 500m x 500m	403	2	0.3(1)	20K	34[2]
Zhang et al. (2012) [86]	Potential field	Room 16m x 10m	–	–	–	600	–
Jian et al. (2014) [57]	Potential field	Corridor with 90° corner	–	–	–	100-200	–
Hoogendoorn et al. (2015) [58]	Potential field	Room 30m x 30m	–	–	–	100-300	–
Jiang et al. (2020) [59]	Potential field	Room 18m x 18m, 1 obstacle 3m x 3m	–	–	–	100-500	–
Sun et al. (2021) [60]	Potential field	Complex office 25m x 30m	–	–	–	100-500	–

- **Macroscopic models**: aim to represent macroscopic characteristics of crowds, such as density, flow, and overall collective dynamics.
- **Hybrid models**: typically couple both microscopic and macroscopic models into one framework to leverage the strengths of each approach.

[1] 25 ms on 2000 agents.
[2] without path search.

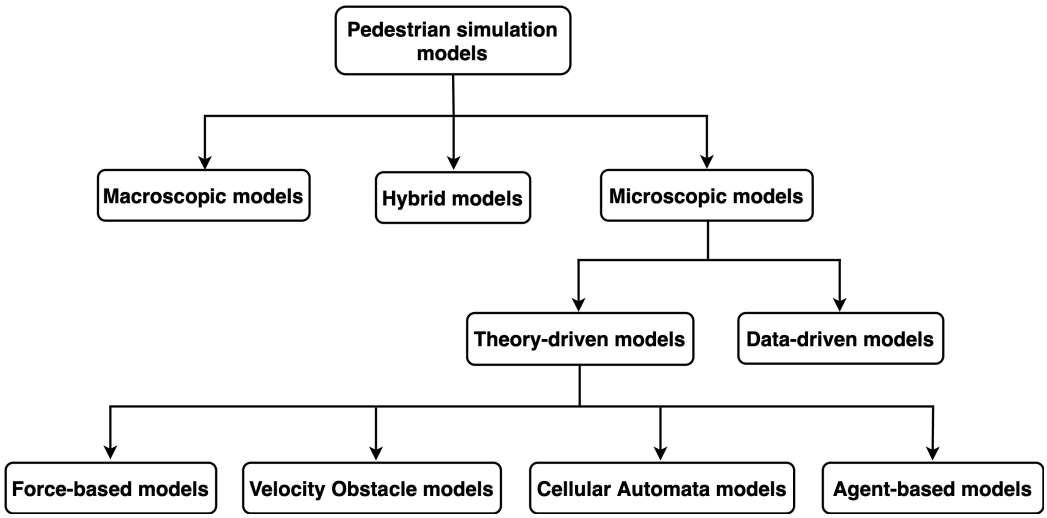


Figure 2.8: Hierarchical classification of operational-level models.

Applying operational-level models for simulating high-density crowds introduces significant questions compared to low-density scenarios:

- One open question regarding whether the performance of these operational-level models, which have proven to be effective in low-density scenarios, can be maintained in high-density contexts.
- The next challenge is how to adapt these operational-level models to capture more realistically emergent pedestrian behaviors and dense crowd phenomena and characteristics in such conditions.
- Finally, a critical concern of how these models are calibrated and validated using high-density crowd datasets to ensure the reliability of their simulation results.

This section seeks answers to these aforementioned questions by first examining the classical model of each approach. It then investigates the applications of these models to highlight key issues likely to arise when applying these models for simulating high-density scenarios. Lastly, the section explores how the classical models have been adapted to better simulate pedestrian behaviors in high-density situations.

Force-based models

The Social Force Model (SFM), proposed by Helbing et al. [13, 61], is one of the most influential models used to simulate pedestrian dynamics. The model is based on the assumption that

pedestrian movement is governed by a set of "social forces". These forces are derived from both intrinsic motivations, such as an individual's destination, and extrinsic factors, including attractive and repulsive effects exerted by neighboring pedestrians and obstacles. The SFM has been proven to replicate realistically well-known crowd phenomena such as lane formation [61] or semi-circle at the bottlenecks [93]. Mathematically, the SFM is grounded in Newton's second law of motion:

$$\mathrm{m}_i \frac{\mathrm{d}\mathbf{v}_i}{\mathrm{d}t} = \mathrm{m}_i \frac{\mathrm{v}_i^0(t)\mathbf{e}_i^0(t) - \mathbf{v}_i(t)}{\tau_i} + \sum_{j \neq i} \mathbf{f}_{ij}(t) + \sum_{W} \mathbf{f}_{iW}(t) \tag{2.1}$$

where m_i represents the mass, $\mathrm{v}_i^0(t)$, $\mathbf{e}_i^0(t)$, and $\mathbf{v}_i(t)$ denote the preferred speed, desired direction vector, and velocity vector of pedestrian i at time t, respectively. In Equation 2.1, the first component characterizes the motivation of pedestrians to move towards their destination at a preferred speed through the acceleration from their current velocity to the preferred velocity over a reaction time of τ_i. The second and third terms, $\mathbf{f}_{ij}(t)$ and $\mathbf{f}_{iW}(t)$, describe the interaction forces with neighboring pedestrians and obstacles (in a given radius), respectively, with the primary focus on addressing collision avoidance behavior.

The emergent behavior of pedestrians in high-density scenarios can be modeled by incorporating additional types of forces into Equation 2.1. For instance, Hebling et al. [13] introduced sliding friction and counteracting body compression forces to address physical interactions between pedestrians in crowded situations. Nuria et al. [18] suggested a pushing force to simulate the pushing behavior of individuals in a panic state. Yu and Johansson [37] proposed an improvement to the repulsion force to more accurately simulate the intense reactions of pedestrians under critical density conditions. The modified repulsion force was specifically designed to exhibit the sudden and rapidly varying changes in forces, thereby generating crowd turbulence phenomena by spreading localized force changes to a larger area. Mohamed et al. [94] developed new forms of forces for the interaction of pedestrians during the mass-gathering Hajj-Tawaf pilgrimage. Karamouzas et al. [95] proposed an anticipatory force based on an implicit integration of a velocity-based energy function. Xu et al. [96] combined a forced-based model and an emotional contagion model to simulate emotional spreading in high-density crowds. Subramanian et al. [97] suggested a modified SFM to model the dynamics of pedestrian groups in a chain-like formulation when navigating in densely populated environments. Likewise, Song et al. [98] proposed an improved SFM that integrates a forward-learning force to address the domino-like phenomenon when pedestrians have physical interactions with each other in a queue.

However, the original SFM exhibits several limitations when applied in high-density scenarios. One major issue is that the repulsion force amplifies significantly as the interpersonal distances between pedestrians decrease. In extremely dense conditions, when these distances become minimal, the repulsion forces can reach very high magnitudes. This leads to abrupt and significant changes in displacement, causing unrealistic oscillatory behaviors [99]. Moreover, the conventional values assigned to the parameters of these forces typically result in pedestrians maintaining a specific distance from other pedestrians and obstacles, thus making it difficult to achieve an extremely high crowd density. Therefore, the parameters of SFM need to be calibrated for the simulation of high-density scenarios. Numerous works have been conducted to calibrate the parameters of SFM for more realistic simulations of dense crowds. Moonsoo et al. [100] employed a maximum likelihood estimation method to calibrate the SFM's parameters using aggregated crowd video data of different density scenarios. Their findings indicated that the influence of forces driving individuals toward their destination decreases significantly as the density increases. In similar research proposed by Haghani and Sarvi [101], the parameters of SFM were calibrated by a multi-directional search algorithm for realistic simulation of pedestrian flows at narrow exits. Their study pointed out that the relaxation time and friction force parameters are the most significant factors in generating accurate pedestrian outflows. Shuaib [102] introduced a 6-order polynomial function of local density to estimate the value of anisotropic angular parameter in [13]. Similarly, Narang et al. [103] suggested density-dependent filters to compute the desired speed in Equation 2.1, with the aim of fitting the simulated velocity with the empirical fundamental diagram [104, 105, 31, 106, 107, 108, 109]. Sticco et al. [110] adjusted the friction coefficient in [13] to regulate the pedestrian flow in extremely congested areas.

Several other improvements have been made to enhance the realism of SFM for dense-crowd simulations. One such improvement was the introduction of a field of vision for pedestrians [18, 94, 111] to gather information about neighboring individuals that they predominantly interact with. Nuria et al. [18] mitigated the unrealistic oscillatory behavior exhibited by the SFM by introducing stopping rules. These rules apply when pedestrians are not in a panic state and are experiencing significant repulsion forces from neighbors moving in the opposite direction in their intended path. Additionally, more accurate representations of the human body have been introduced, such as spheropolygons [112], three-circle shape [113], and elliptical shape [114], to represent better heterogeneous characteristics of individuals in the crowds and the high-density situations.

Velocity Obstacle models

The Velocity Obstacle (VO) model was introduced by Paolo Fiorini and Zvi Shiller in 1993 [115] for the avoidance maneuver strategy of robots in dynamic environments. The VO model constructs a geometric space, denoted as VO_B^A, which represents the set of velocities that would result in a collision with an obstacle (B that can be another robot or an obstacle from the environment) [116]:

$$\mathrm{VO}_B^A(\mathbf{v}_B) = \{\mathbf{v} \mid \lambda(\mathbf{p}_A, \mathbf{v} - \mathbf{v}_B) \cap \mathrm{B} \oplus -\mathrm{A} \neq \emptyset\} \tag{2.2}$$

where $\mathbf{p}_A$, $\mathbf{p}_B$ and $\mathbf{v}_A$ and $\mathbf{v}_B$ represent the positions and velocities of two disc-shaped agents A and B, respectively. In addition, $\lambda(\mathbf{p}, \mathbf{v})$ is a ray that originates from point $\mathbf{p}$ and extends in the direction of $\mathbf{v}$:

$$\lambda(\mathbf{p}, \mathbf{v}) = \{\mathbf{p} + t\mathbf{v} \mid t \geq 0\} \tag{2.3}$$

Furthermore, $\mathrm{A} \oplus \mathrm{B}$ represents the Minkowski sum of geometric shapes of agents A and B, whereas $-\mathrm{A}$ denotes the reflection of agent A's geometry:

$$\mathrm{A} \oplus \mathrm{B} = \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in \mathrm{A}, \mathbf{b} \in \mathrm{B}\} \tag{2.4}$$

$$-\mathrm{A} = \{-\mathbf{a} \mid \mathbf{a} \in \mathrm{A}\} \tag{2.5}$$

Agent A is predicted to collide with agent B at a future time if a ray originating from agent A's position $\mathbf{p}_A$ in the direction $\mathbf{v}_A - \mathbf{v}_B$ (the relative velocities of agents A and B) intersects the Minkowski sum of B and $-\mathrm{A}$ located at $\mathbf{p}_B$. As illustrated in Figure 2.9a, $VO_B^A(\mathbf{v}_B)$, called velocity obstacle of A induced by B, is the set of velocities that, if adopted by agent A, would lead to a collision between agents A and B. The collision-free velocity is chosen from the set of feasible velocities such that it lies outside of the union of all velocity obstacles induced by neighboring agents and obstacles.

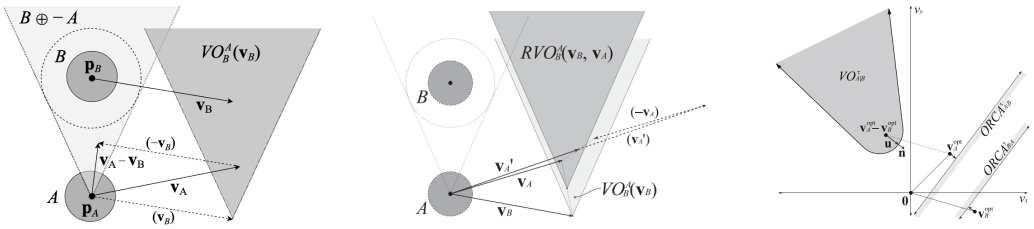
Nevertheless, this avoidance maneuver does not consider the possibility that other agents would adopt new velocities as well; this leads to unrealistic oscillations in the trajectories coming from the reevaluation of other pedestrians' unexpected trajectories (for an in-depth description, refer to Section III. C in [62]). To address this problem, Van den Berg et al. proposed Reciprocal Velocity Obstacle (RVO) [62] (as illustrated Figure 2.9b), which assumes that all moving agents will reciprocally attempt to avoid collision through similar reasoning:

$$\mathrm{RVO}_B^A(\mathbf{v}_B, \mathbf{v}_A) = \{\mathbf{v} \mid 2\mathbf{v} - \mathbf{v}_A \in \mathrm{VO}_B^A(\mathbf{v}_B)\} \tag{2.6}$$

The RVO has demonstrated its effectiveness in producing collision-free and non-oscillatory trajectories for each agent [62]. Building upon the achievement of RVO, Van den Berg et al. proposed a further improvement called Optimal Reciprocal Collision Avoidance (ORCA) [63] (as illustrated in Figure 2.9c) for choosing an optimal velocity that avoids collisions from multiple agents simultaneously:

$$\mathrm{ORCA}^{\tau}_{A|B} = \{\mathbf{v} \mid (\mathbf{v} - (\mathbf{v}_A + \frac{1}{2}\mathbf{u})) \cdot \mathbf{n} \geq 0\} \tag{2.7}$$

where $\mathbf{u}$ is defined as the vector extending from $\mathbf{v}_A - \mathbf{v}_B$ to the closest point on the boundary of velocity obstacle (as depicted in Figure 2.9c). Here, $\mathbf{n}$ denotes the unit vector in the direction $\mathbf{u}$. Geometrically, $\mathbf{u}$ quantifies the minimum necessary adjustment in relative velocity between agents A and B to guarantee collision-free movement over τ seconds.



(a) VO [116] **(b)** RVO [62] **(c)** ORCA [63]

Figure 2.9: Velocity obstacle models [116, 62, 63].

In densely populated environments, the significantly growing number of neighboring agents can result in velocity obstacles encompassing the entire range of feasible velocities. Consequently, identifying a collision-free velocity outside these velocity obstacles requires relaxing some constraints. In such scenarios, RVO [62, 81] enables choosing new velocities within the reciprocal velocity obstacles by employing a penalty function that maximizes the time-to-collision with nearby agents while minimizing the difference with the preferred velocity. Likewise, ORCA [63] determines a velocity that lies in the optimal reciprocal collision avoidance set induced by as many neighbors as possible. Additionally, numerous modifications of velocity obstacle formulations have been proposed for dense crowd simulations, such as truncated velocity obstacle cone [117], hybrid coupling of RVO and VO for oscillation-free navigation [118], or velocity obstacle constructing for agents with elliptical shapes.

Nevertheless, the choice of a new velocity from the feasible velocity set is performed in a way that minimizes its difference from the preferred velocity. As a result, agents still tend to choose velocities with high magnitudes in high-density situations. This leads to an

inconsistency between the simulated and the empirical fundamental diagram (i.e. the relationship between velocity and density) [104, 105, 31, 106, 107, 108, 109]. To handle this issue, a more accurate determination of agent's preferred speeds is performed by incorporating physiological and psychological constraints like density filter [119, 103] or the relationship between the stride-walking speed and the personal space [120].

The velocity obstacle methods predominantly focus on simulating collision avoidance behavior; however, different emergent behaviors of pedestrians are observed in high-density situations mainly due to physical interactions. To integrate fine representation of physical interactions into VO-based methods, Kim et al. [121, 122] integrated force-related constraints into the formulation of velocity obstacles. The details of this method are as follows:

$$\mathbf{v}_A^{\mathrm{f}} = \mathbf{v}_A + \frac{\mathbf{f}_A}{\mathrm{m}_A}\Delta t \tag{2.8}$$

$$\mathrm{FC}_A = \{\mathbf{v} \mid (\mathbf{v} - \mathbf{v}_A^{\mathrm{f}}) \cdot \hat{\mathbf{f}}_A \geq 0\} \tag{2.9}$$

$$\mathrm{PV}_A = \mathrm{FC}_A \cap \bigcap_{B \neq A} \mathrm{ORCA}_{A|B} \tag{2.10}$$

where Δt denotes the simulation time step, m_A and $\mathbf{f}_A$ represent the mass and aggregate force applied to agent A, respectively. The force constraint FC_A is induced by $\mathbf{f}_A$. Geometrically, FC_A defines a half-plane bounded by a line that intersects $\mathbf{v}_A^{\mathrm{f}}$ and is perpendicular to the normalized force $\hat{\mathbf{f}}_A$ of $\mathbf{f}_A$. This half-plane includes velocities whose differences from the current velocity $\mathbf{v}_A$ are greater than or equal to the minimum changes in velocity caused by the aggregate force $\mathbf{f}_A$. A new velocity for agent A is selected from the permitted velocity set PV_A, which is the intersection of FC_A and ORCA_A.

Furthermore, several works have been developed to integrate different models with VO-based models for a more diverse representation of pedestrian behavior levels. For instance, [122, 123] proposed a combination of Finite State Machine (FSM) with RVO to simulate high-level decision-making pedestrian behaviors at the Hajj-Tawaf pilgrimage mass-gathering event. The FSM includes states that represent different behaviors, and each state is activated based on temporal and spatial conditions as well as the agent's state. Initially, the FSM evaluates and adjusts parameters associated with each behavior for the agents before performing local collision avoidance behavior using RVO (as shown in Figure 2.10). Similarly, Sudkhot et al. [124] introduced a framework for faster and smoother navigation of agents by combining the Belief-Desire-Intention (BDI) [125, 126] model for modeling the high-level path planning behavior and the RVO for low-level collision avoidance behavior. Golas et al. [127] modeled the density-sensitive behavior of pedestrians by incorporating long-range collision anticipation. As a

result, agents are capable of predicting congested areas in advance and adjusting their trajectories accordingly. By steering away from these congested areas earlier, agents can maintain smoother trajectories and avoid abrupt changes in direction.

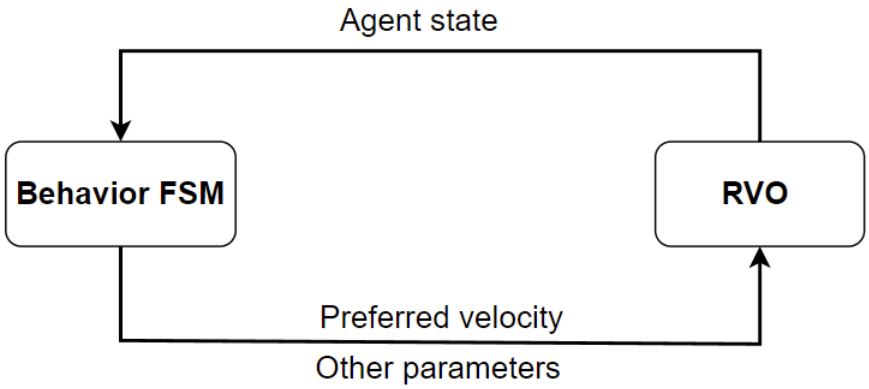


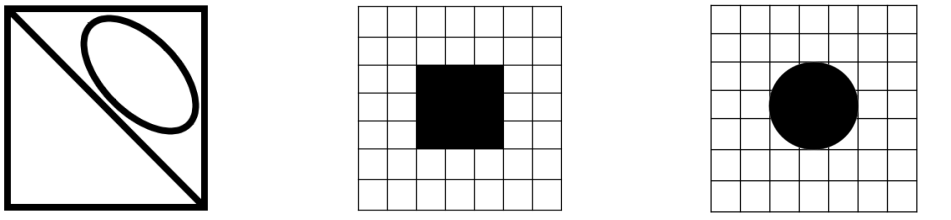
Figure 2.10: Combination of FSM and RVO in [123].

Cellular Automaton models

Initial applications of Cellular Automaton (CA) models for simulating pedestrians were proposed by Burstedde et al. [12] in 2001 and Kirchner and Andreas [38] in 2002. These models consider the environment as a grid divided into cells. Each cell typically has a size of $40cm \times 40cm$, with the aim that each cell can hold a maximum of one pedestrian. The movement of pedestrians at each simulation step is determined by the transition probability computed using the concept of *floor field*. This floor field can be understood as an implicit additional grid of cells beneath the pedestrian-occupied grid [12]. It consists of two elements: *static floor field* and *dynamic floor field*. The static floor field, which is determined by the distance between cells and the destination, provides a directional guide for pedestrian toward their targets. In contrast, the dynamic floor field, inspired by the concept of chemotaxis, describes the virtual trail left behind by pedestrians as they move and how it impacts other pedestrians move. This virtual footprint can diffuse or decay over time. Both static and dynamic floor fields contribute to the calculation of the transition matrix, which defines the probability of moving to adjacent, unoccupied cells (using a Moore neighbor, i.e. with 8 neighbors cell). The CA models have been demonstrated to produce well-known crowd phenomena such as lane formation in the bidirectional corridor or a semi-circle at the bottleneck [12, 38].

However, the standard cell size of $40cm \times 40cm$ used in the CA models limits the maximum

density to 6.25 $\mathrm{ped}/\mathrm{m}^2$, making it challenging to represent extremely high densities of 8–9 $\mathrm{ped}/\mathrm{m}^2$ observed in real-world scenarios [31]. Furthermore, while simplifying a pedestrian to an individual cell can avoid local collisions due to its nature, it restricts the direction and speed to limited discrete values at any given time. The traditional CA models [12, 38] employed either the Moore or the Von Neumann neighborhood, which constrains pedestrian movements to four or eight possible directions. To address these limitations, finer discretizations of the environment have been proposed to allow the simulation of higher densities and more flexible pedestrian movement. In these cases, smaller cell sizes are applied, and consequently, the pedestrian shape covers multiple cells, such as triangular cell [128], group of 3×3 cells [129, 130], group of 4×4 cells [131], group of 20 cells [132], group of 39 cells [133] (as shown in Figure 2.11).



(a) Jingwei et al. [128] **(b)** Bandini et al. [129] **(c)** Guo and Huang [130]

(d) Zhijian et al. [131] **(e)** Bazior et al. [132] **(f)** Sarmady et al. [133]

Figure 2.11: Finer representations for pedestrian shape.

In addition, representing pedestrian shapes by multiple smaller cells leads to a broader range of movement possibilities. Several studies have expanded the directional choices to facilitate smoother pedestrian movements. For example, Jingwei et al. [128] implemented triangular cells for the representation of pedestrian shapes and enabled pedestrian movement in each cell to 14 possible directions corresponding to its 14 neighboring cells. Claudio et al. [134] added a position at the center of each cell edge and allowed pedestrians to move to the midpoint of the next cell edge if pedestrians counted in that cell less than a certain threshold. An extended Moore neighborhood was suggested by Huo et al. [135] to simulate the stampede behavior of pedestrians during evacuation scenarios in which pedestrians were enabled to walk two cells in

one simulation step. Zhang et al. [136] addressed the ability of CA models to simulate group behavior by developing transition rules for pedestrian movement to simulate pedestrians walking in pairs with different decision-making capabilities at the metro station.

In the CA models, the emergent behavior of pedestrians in crowded situations can be modeled by modifying the floor fields to address physical interactions [137, 130], group behaviors [138], or by introducing new types of floor field such as force floor field [139, 140] for physical interactions, local view floor field [141] for replanning behavior, anticipation floor field [142, 134] for predicting future collisions, interplay floor field [143] for maintaining social distance among agents, and environmental floor fields such as wall floor field [134, 143] or water flood field [144]. These various floor fields influence the calculation of the transition matrix, which in turn specifies the probability of choosing neighboring cells for movement in the next simulation step.

Agent-based models

Agent-based modeling (ABM) is a bottom-up approach in which agents are considered unique and independent entities capable of interacting locally with other individuals and their environment [145]. In crowd simulation, an agent is a rule-based entity representing an individual pedestrian. The agent-based approach provides the ability to simulate heterogeneous crowds and can be combined with various other models [146]. Despite the flexibility in modeling pedestrian behavior, this method requires careful consideration of the trade-off between model complexity and computational costs.

It is worth acknowledging that the definitions of agent-based methods can sometimes be ambiguous in the literature. Therefore, methods such as SFM, CA, and VO can occasionally be categorized as part of the agent-based approach due to their consideration of pedestrians as discrete individual entities rather than aggregated flows. In this manuscript, I limit the agent-based models to models in which pedestrian agents exhibit autonomous decision-making capabilities. Specifically, agents in ABMs typically possess advanced cognitive processes that allow them to sense and evaluate their environment. Their actions then adapt based on these perceptions, rather than following strictly predetermined rules. For example, in the scenario of people evacuating from a building, pedestrians might choose the nearest exit in low-density situations, but in high-density cases, they might opt for the quickest exit or follow a leader.

Shao and Terzopoulos [64] developed a comprehensive framework that integrates perceptual, behavioral, and cognitive components to simulate complex decision-making behaviors of pedestrians in various environments. Zhou et al. [147] embedded three different fuzzy inference

systems into the decision-making behaviors of pedestrians, each inference system corresponding to a specific level of pedestrian behavior [29]. These systems utilized information extracted from agents' perceptions, prior experience, and knowledge as inputs. The outputs from these inferences were then combined to compute the turning angle and speed for pedestrian movement in the next simulation step. Sharma et al. [148] trained a neural network to predict the adaptive behavior of pedestrians during emergency evacuation. Their study also incorporated a fuzzy logic system to infer pedestrians' speed based on their emotional state and stress levels.

Moreover, agents have the capability to gather information from their neighboring areas, which can greatly influence their immediate decisions and actions, especially in high-density conditions. Research by Zheng et al. [108] has demonstrated that in densely populated areas, the impact of neighbors ahead on pedestrian speed and acceleration is much stronger than that from lateral neighbors. The traditional approaches [13, 62, 63] consider all neighboring pedestrians within a predefined distance for local interactions. While this method is advantageous for its simplicity and computational efficiency, more realistic simulations can be achieved in crowded conditions by disregarding non-interacting neighbors. Several works have been proposed to filter only neighbors that pedestrians actually interact with, by developing visual fields such as a rectangle [18] or a cone shape [64, 147, 149, 150].

In high-density environments, pedestrians often cannot maintain their preferred walking speed due to limited personal space. This space restriction can lead to psychological discomfort, which in turn can affect decision-making processes and result in different emergent behaviors in crowded environments [26]. Leader-follower behavior is one of the typical emergent behaviors in these scenarios. This behavior involves the tendency of pedestrians to follow nearby people ahead who are moving in similar directions to reduce the effort during navigation through dense crowds [151]. Samuel et al. [65] carried out an experiment in one-dimensional pedestrian traffic to investigate and model the following behavior of pedestrians. In their model, agents accelerate their speed based on the relative movement characteristics of the agent in front of them. Godoy et al. [151] computed the next velocity of the following behavior based on data collected from the most similar and most constrained agents. Liang et al. [152] classified pedestrian crowds into different groups based on their similarity of targets and velocities. The leader of each group is selected based on proximity to the target and inability to follow another member. The remaining members follow the leader while maintaining collision-free interactions with inter-group members. Similarly, Liu et al. [153] suggested different criteria for leader selection during emergency evacuations, including environmental familiarity, distance to exits, and navigation knowledge. Mingbi et al. [19] proposed a role-dependent model identifying leaders as those with current speeds much higher than the group average speed. A double-layer decision

model was suggested by Li et al. [154] to address pedestrian decisions in choosing a detour when perceiving a congested area in front. Xie et al. [155] developed an information-theoretic method to model the spontaneous leader-follower behavior of evacuees during crowd emergency scenarios.

Data-driven models

Over the last decades, the exponential growth in data availability coupled with significant improvements in computational capabilities have made neural network methods a powerful tool for investigating hidden patterns in large datasets. These approaches take inspiration from the interconnected neural network structures found in the human brain. Neural network models have been proven effective in solving complex problems and are employed in various applications such as computer vision, autonomous vehicles, and business analytics.

Neural network methods for predicting human trajectories present a different approach, with different objectives, compared to knowledge-based models for pedestrian simulation. Knowledge-based models are established on predetermined rules predominantly designed to model pedestrian behaviors, with trajectory predictions generated as the secondary outcome [28]. On the other hand, neural network methods employ data-driven learning mechanisms to learn underlying patterns from pedestrian datasets and primarily focus on predicting pedestrian trajectories based on these learned representations.

The neural network models employ historical positions of pedestrians for a given number of time steps as inputs[3], in conjunction with supplementary neighboring information about the environment and nearby pedestrians, to predict the future trajectory of a target individual for a prediction time horizon. One of the most well-known contributions to this field is the Social-LSTM (Long Short-Term Memory) developed by Alahi et al. [11]. The Social-LSTM incorporates the social pooling mechanism to aggregate hidden states of neighboring individuals in a scene, thereby taking local interactions into consideration during prediction. Another influential data-driven model for predicting human trajectory is the Social-GAN (Generative Adversarial Networks) suggested by Agrim et al. [66]. This Social-GAN architecture comprises an RNN (Recurrent Neural Network) Encoder-Decoder generator coupled with an RNN-based encoder discriminator that considers the influence of all pedestrians within the scene. Various adaptations of LSTM-based and GAN-based models have been proposed to enhance prediction accuracy in diverse environments, such as DESIRE framework [157], Social Attention model [158], Sophie

[3]Trajectory prediction is thus very similar to the problem of prediction using time series, for which LSTM model has been developed [156].

model [159], DH-SARL model [160]. Neural network models have demonstrated significant effectiveness in predicting pedestrian trajectories in low-density scenarios, with predictions showing better distance accuracy than knowledge-based models in such conditions [10].

Nevertheless, these models still encounter difficulties when applied in high-density conditions. The primary goal of neural network models is to minimize the distance error compared to actual trajectories, which explains why distance-based metrics such as average displacement error (ADE) and final displacement error (FDE) are frequently used as training functions and evaluation criteria. When it comes to higher densities, studies have shown that training and evaluation should not rely only on distance errors [10] [161]: a training function only focusing on the distance-based error may still generate unrealistic phenomena such as a high number of collisions in predicted trajectories [7]. To tackle this problem, different solutions have been proposed to design better training functions for high-density situations. These include integrating additional loss components for collisions into the training function, such as time-to-collision loss [9] or interaction energy loss [7], alongside traditional distance loss terms. Another consideration for improving neural network models in crowded scenarios is the prediction time horizon. Current models typically predict pedestrian trajectories for 3.0 – 5.0 seconds in the future [28]. However, this time frame may need to be re-evaluated for high-density scenarios, as the number of interactions increases significantly in these situations. Lastly, neural networks are typically trained on low-density datasets due to the limited availability of high-quality, high-density datasets.

Macroscopic models

Macroscopic models are a category of models that consider the system from a macroscopic point of view, attempting to model the evolution of macroscopic numbers such as average speed, and density..., instead of considering individual pedestrians. Classical modeling approaches include partial differential equations to simulate crowd dynamics. These models consider crowd movement as a continuous flow and focus on addressing overall crowd behavior and movement while disregarding the modeling of detailed individual-level interactions.

One widely used technique for simulating crowd flow is treating it as analogous to fluid dynamics. This approach applies the principle of conservation of flow through the continuity equation [162]:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{2.11}$$

where ρ and $\mathbf{v}$ represent the density and velocity vector of the crowd flow, respectively. This fundamental equation has been found in numerous macroscopic models for simulating crowd

dynamics [67, 163, 164]. Building upon this foundation, Hughes introduced a comprehensive continuum theory for pedestrian flows [14, 165] that integrates both physical and psychological characteristics of individuals into the modeling framework:

$$\frac{\partial \rho}{\partial t} - \nabla \cdot (\rho g(\rho) f^2(\rho) \nabla \phi) = 0 \tag{2.12}$$

where $f(\rho),\ g(\rho)$ denote density-dependent coefficients that affect pedestrian walking speed and level of discomfort, respectively, whereas $\nabla\phi$ represents the principal direction of pedestrian movement at any given point in the environment. To cap the density in simulations, Narain et al. [67] proposed a unilateral incompressibility condition for the density variable in Equation 2.11. This condition ensures that the density cannot exceed a certain threshold, thus preventing further compression when maximum density is reached:

$$\rho \leq \rho_{\max} = \frac{2\alpha}{\sqrt{3} d_{\min}^2} \tag{2.13}$$

where $\rho_{\max}$ denotes the maximum possible density, which is determined by two key parameters: a constant factor α and the minimum interpersonal distance d_{min} between individuals in the crowd. Furthermore, additional components can also be incorporated to account for complex dynamics observed in crowd movements. For instance, Jiang et al. [163, 164] proposed supplementary equations that introduce relaxation and anticipation terms to simulate the traffic instability phenomenon emerging in dense crowds:

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{v} + c^2(\rho)\frac{\nabla \rho}{\rho} = \frac{U_e(\rho)\vec{v} - \mathbf{v}}{\tau} \tag{2.14}$$

where $c, \tau, U_e, \vec{v}$ represent the sonic speed, relaxation time, equilibrium value, and desired direction, respectively.

In contrast, Treuille et al. [15] introduced a continuum model where global navigation is governed through a dynamic potential field derived from the eikonal equation (Equation 2.15):

$$\|\nabla \phi(\mathbf{x})\| = \mathrm{C} \tag{2.15}$$

$$\mathbf{v} = -f(\mathbf{x}, \theta)\frac{\nabla \phi(\mathbf{x})}{\|\nabla \phi(\mathbf{x})\|} \tag{2.16}$$

where ϕ denote the potential, and C function and unit cost associated with the direction $\nabla\phi$. The cost of reaching the destinations is estimated based on various factors, including distance, time, and discomfort. Pedestrian movement is characterized by a velocity vector that points opposite

to the gradient of the function in Equation 2.15. The magnitude of this velocity is determined by evaluating the speed field $f(\mathbf{x}, \theta)$ at the location $\mathbf{x}$ with moving direction θ. A similar study by Jiang et al. [163] aimed to minimize the total instantaneous walking cost from start to target points by considering both instantaneous equilibrium travel cost and discomfort field.

In recent developments, crowd simulation models [166, 167, 168] inspired by Smoothed Particle Hydrodynamics [169] have been proposed to model crowd dynamics at extremely high-density conditions. Additionally, macroscopic models are frequently used to simulate pedestrian crowds in high-density scenarios because these models assume crowd flows as a continuum flow, which is particularly valid when density is high. Macroscopic models have demonstrated their efficiency in simulating macroscopic crowd phenomena under extremely high-density conditions like shockwave propagation [170, 166, 168, 171, 172], panic propagation [173, 174, 175, 176], crowding force [170, 174]. Moreover, by focusing on aggregate behaviors, these models can also effectively simulate dense crowds at a large scale with reduced computational cost compared to microscopic models.

Hybrid models

Hybrid models typically combine microscopic models and macroscopic models into a single framework to leverage the advantages of both approaches [177]. This class of modeling incorporates the detailed resolution of microscopic models and takes the computational efficiency of macroscopic models, thus offering a balance between the two. Hybrid models are particularly advantageous in contexts requiring an in-depth analysis of the crowd-level phenomena alongside specific individual behaviors and interactions. A fundamental challenge of these approaches is the need for effective communication and consistent transition between macro and micro scales. This involves aggregating individual-level data into a continuous framework for macro-level simulations, as well as disaggregating macro-level data (like density distribution) back into micro-level simulations.

Hybrid models can be classified into two main types: region-based models and auto-switch models. Region-based models [3, 69, 68, 178, 70, 179, 180] partition the environment into separate regions. Each region is simulated using either a microscopic or macroscopic model and all regions are simulated concurrently. The selection of a suitable simulation model for each region depends on expert knowledge and the current state of the simulation. For instance, regions expected to be crowded can be simulated by macroscopic models, whereas microscopic models are used for regions predicted to have a lower density. Furthermore, the dynamic nature of crowd movement leads to changes in crowd states over time. Thus, there is the need for

the ability to adjust simulation accordingly. This can be achieved by dynamically switching simulation models for each area using specific conditions like density levels [179, 180]. In these cases, the criteria for transition and data synchronization mechanisms are triggered when agents move from one area to a new region. Moreover, boundary cells [68, 179, 180], which are additional cells positioned between microscopic and macroscopic simulation regions, are normally applied to enhance the consistency of data conversion and synchronization.

Conversely, auto-switch models [181, 182] combine microscopic and macroscopic models in a way that these models can be switched at runtime. In these frameworks, only one model is executed at a time during the simulation. The switch between models is defined by predetermined rules, typically based on specific events or crowd characteristics observed in the simulated environment.

2.4.4 Comparison of operational level models

This section proposes a set of critical criteria for evaluating and comparing different operational-level models in high-density scenarios. The first criterion investigates the models' ability to model emergent behaviors and phenomena typically observed in densely populated environments. This includes the capacity to simulate physical interactions, group behavior, following behavior, and replan behavior. The second criterion focuses on quantitative performance metrics, including time step duration, maximum simulated density, number of simulated agents, and computational efficiency. The third criterion assesses the validation of operational-level models against the empirical Fundamental Diagram [104, 105, 31, 106, 107, 108, 109] and real-world crowd data. Finally, the fourth criterion examines multi-level modeling capabilities, which specifically evaluate how effectively the operational-level model can integrate behaviors at strategic and tactical levels.

Table 2.3 presents the specifications of all the operational-level models from the selected corpus based on the aforementioned criteria. The following syntax is used to indicate the ability to simulate each characteristic:

- ✓: The model is able to address the specified characteristic.
- $\times$: The model is unable to simulate the specified characteristic.
- $-$: Unknown.

Data from Table 2.3 is analyzed in the next sections to compare and evaluate the operational-level models.

Behavior capability

Realistic simulations of dense crowds necessitate the accurate modeling of emergent behaviors and phenomena. In these scenarios, individuals display a broader range of behaviors compared to low-density situations, which include physical interactions, group behavior, following behavior, and replan behavior in response to the surrounding environment. Such emergent behaviors contribute a significant impact on overall crowd dynamics as well as potentially increase risks in high-density settings.

The data presented in Table 2.3 reveals that the SFMs have been widely selected to model diverse emergent behaviors of pedestrians. Additionally, the classical SFM is the only model that shows the capacity to model emergent behaviors in high-density situations such as physical interactions, whereas other classical models need necessary modifications to represent these behaviors.

In crowded scenarios, simulation models are modified to represent the emergent behaviors of pedestrians. In SFMs, modifications take the form of incorporating additional terms into Equation 2.1 that describe new forces. However, extreme densities can induce strong and rapidly varying forces, which can lead to unrealistic behavior such as oscillatory movement [99]. Therefore, calibrating and validating parameters using real-world data is a crucial next step to ensure the realism of these models and their applicability to specific scenarios.

In CA models, the cell size must be decreased to address physical interactions among pedestrians in dense situations. This is because the traditional cell size of 40cm x 40 cm [12] was chosen based on an approximation of space occupied by a single pedestrian, which inadvertently prevents physical contact between pedestrians in the model. Furthermore, when introducing new floor fields to represent new behavior, it is crucial to concurrently modify the transition matrix to account for the impact of these fields on pedestrian movement. However, integrating new floor fields may increase model complexity and require substantial alterations in algorithm implementation.

In ABMs, modeling emergent behaviors in high-density environments relates to introducing variables that represent decision-making behaviors. This method requires a prior determination of emergent behaviors to be modeled and can lead to significant computational demands when dealing with a large number of agents with complex decision-making behaviors.

Contrarily, VO models and macroscopic approaches have limited capabilities in simulating emergent behaviors of pedestrians in densely populated scenarios. VO models require the establishment of a new velocity obstacle space that takes into account velocity constraints for new behaviors. Macroscopic models primarily focus on addressing the macroscopic characteristics

of crowds rather than individual behaviors. In contrast, the capacity of hybrid models to simulate emergent behaviors is heavily dependent on specific submodels integrated into the hybrid framework.

Current data-driven models show significant limitations in modeling emergent behaviors of pedestrians in crowded environments. Existing deep learning architectures are predominantly trained based on low-density datasets using distance-based loss functions, which is not sufficient in high-density situations. Thus, there is a need for further improvements to make these models applicable in high-density scenarios. These improvements may involve developing new neural network architectures or modifying existing ones to better accommodate high-density scenarios. Furthermore, data-driven models encounter several challenges regarding the interpretability of predictions and the need for more availability of high-density pedestrian datasets.

Model performance

Various quantitative metrics are used to evaluate the performance of operational-level models, including the ability to simulate with different time step values, maximum simulated density, number of simulated agents, and computational efficiency. Models exhibiting minor sensitivity to variations in simulation time steps tend to generate more reliable and stable simulation outcomes. Furthermore, smaller simulation time steps can result in smoother trajectories, but can also increase computational costs. Next, the maximum simulated density serves as a criterion to evaluate the model's capacity to simulate high-density crowds. Lastly, the computational efficiency and scalability of the operational-level models are assessed by analyzing the correlation between computational time and the number of simulated agents.

As shown in Table 2.3, SFMs, ABMs, and macroscopic models typically employ small simulation time steps, ranging from $0.01s$ to $0.1s$, to maintain stability. Conversely, VO models can accommodate a wider range of simulation time steps (0.01 - $0.25s$). With CA and data-driven models, the simulation time steps are primarily influenced by the cell size and the frame rate of training datasets to ensure an accurate representation of pedestrian speed, generally around $0.3s$ and $0.4s$, respectively. With hybrid models, the primary and still open challenge is synchronizing simulation time steps between different submodels within the hybrid framework.

Regarding simulated crowd density, as presented in Table 2.3, SFMs, VO models, and ABMs have commonly been utilized to address densities ranging from 2 to 6 $\mathrm{ped}/\mathrm{m}^2$. Under extremely high-density scenarios, these models may exhibit unrealistic phenomena like vibrations in trajectories [18] or stuck phenomena [19]. Therefore, the parameters of these models must be carefully calibrated to simulate realistic pedestrian behaviors accurately in such extreme

conditions. Conversely, CA models and macroscopic models demonstrate the capacity to simulate extreme densities of 6 - 8 $\mathrm{ped}/\mathrm{m}^2$. The ability of CA models to handle such densities is based on their intrinsic attributes, which allow more effective spatial utilization through the discretization of the environment into small cells. Furthermore, CA approaches disregard the concept of personal space as well as inter-agent repulsion in their models. As a result, these models can simulate density up to the theoretical maximum, provided that every cell is occupied by pedestrians.

In terms of computational efficiency and scalability, SFMs, CA models, and ABMs are generally used for small-scale simulations of 100 - 500 agents and medium-scale simulations of 500 - 5K agents. In contrast, VO, macroscopic, and hybrid models exhibit the capacity to simulate large-scale crowds of 5K - 100K agents at reasonable computational costs. Currently, the prediction of data-driven models is predominantly focused on the trajectory of a single primary pedestrian.

Validation

Validating operational-level models is crucial for verifying the reliability and accuracy of simulation results. The validation step includes a comparative analysis of simulation outcomes with empirical fundamental diagrams [104, 105, 31, 106, 107, 108, 109] and other data obtained from real-world observations. Fundamental diagrams serve as theoretical benchmarks for comparing general crowd movement patterns, as they provide fundamental relationships among average density, speed, and flow of crowds. On the other hand, validating operational-level models using other real-world data, such as trajectory or contact forces, enables more targeted validation for scenarios with specific behavior and phenomena.

Table 2.3 indicates that ABMs and data-driven models are validated more often than other models due to their regular integration of empirical data for calibrating model parameters. On the other hand, macroscopic models and hybrid models exhibit a notable lack of validation in the literature. The classical models of SFMs, VO models, and CA models have historically been under-validated against empirical fundamental diagrams and additional real-world data. However, recent research has started to handle this shortfall to enhance the realism and reliability of high-density crowd simulation. This trend highlights the increasing significance of the validation step in verifying the accuracy of operational-level models for high-density contexts.

Multi-level modeling

The integration of operational-level models with higher-level modeling at strategic and tactical levels plays an essential role in developing a comprehensive framework that captures holistically pedestrian behaviors. This integration improves the adaptability of operational-level models for effectively simulating various environmental contexts and situations.

CA models and macroscopic models intrinsically include tactical-level modeling in their models via navigation mechanisms like floor fields or potential fields. Therefore, these models are less adaptable for integrating novel tactical-level models, such as graph networks or navigation mesh. In contrast, SFMs, VO models, ABMs, and hybrid models demonstrate greater flexibility in incorporating strategic and tactical models due to the communication between these levels through preferred targets and preferred velocities, which are then directly used as parameters of operational-level models. Finally, existing data-driven models encounter difficulties in integrating higher-level modeling due to their primary focus on predicting short-term future trajectories.

Table 2.3: Comparison results of operational level models based on various evaluation criteria (PI: physical interaction, GB: group behavior, FB: following behavior, RB: replan behavior, TS: time step (s), MD: max density ($\mathrm{ped}/\mathrm{m}^2$), NA: number of agents, RT: run time (fps), S: strategic, T: tactical, O: operational).

Articles	**Class**	**Behavior capability**				**Model performance**				**Validation**		**Multi-level**		
		PI	GB	FB	RB	TS	MD	NA	RT	Fit FD	Real data	S	T	O
Helbing et al. (2000) [13]	SFM	✓	✓	×	×	–	–	200	–	–	–	×	×	✓
Lakoba and Kaup (2005) [99]	SFM	✓	✓	×	×	<0.0625	–	100	–	–	–	×	×	✓
Yu and Johansson (2007) [37]	SFM	✓	✓	×	×	–	9	–	–	–	–	×	×	✓
Nuria et al. (2007) [18]	SFM	✓	✓	✓	✓	–	–	1800	25	–	–	×	✓	✓
Guy et al. (2010) [91]	SFM	×	×	×	✓	–	4.0	10K	15.1	–	✓	×	✓	✓
Moonsoo et al. (2013) [100]	SFM	×	×	✓	×	–	–	–	–	–	✓	×	×	✓
Shuaib (2014) [102]	SFM	✓	✓	×	×	0.001	4.5	100-400	–	✓	–	×	×	✓
Best et al. (2014) [119]	SFM	×	×	×	×	0.0625	3.96	–	–	✓	✓	×	✓	✓
Mohamed et al.(2015) [94]	SFM	✓	✓	✓	×	–	8	–	–	✓	–	×	✓	✓
Karamouzas et al. (2017) [95]	SFM	×	✓	×	×	0.01-0.4	–	5K	4.2	–	–	×	✓	✓

Jingni et al. (2019) [113]	SFM	✓	×	×	×	–	8.6	–	–	✓	✓	×	×	✓
Haghani and Sarvi (2019) [101]	SFM	✓	✓	×	×	–	3	200	–	–	✓	×	×	✓
Sticco et al. (2020) [110]	SFM	✓	×	×	×	0.05	9	50-400	–	✓	–	×	×	✓
Xu et al. (2020) [96]	SFM	✓	×	×	×	–	–	60-250	–	–	–	×	×	✓
Kolivand et al. (2021) [183]	SFM	✓	✓	✓	×	–	–	4500	–	–	✓	×	×	✓
Wang et al. (2022) [111]	SFM	✓	✓	×	✓	–	≈ 5.3	–	–	✓	–	×	✓	✓
Subramanian et al. (2022) [97]	SFM	✓	✓	✓	×	–	–	–	–	✓	–	×	×	✓
Zanlungo et al. (2023) [114]	SFM	✓	✓	×	×	0.05	3	–	–	–	✓	×	×	✓
Song et al. (2023) [98]	SFM	✓	×	×	×	0.01	9	–	–	–	✓	×	×	✓
Van den Berg et al. (2008) [62]	VO	×	×	×	×	0.25	–	1000	12.5	–	–	×	×	✓
Van den Berg et al. (2008) [81]	VO	×	×	×	×	0.25	–	20K	2(6)[4]	–	–	×	✓	✓
Guy et al. (2009) [117]	VO	×	×	×	×	–	–	10K	302(32)	–	–	×	✓	✓
Van den Berg et al. (2011) [63]	VO	×	×	×	×	0.25	–	5K	42.4(8)	–	–	×	×	✓
Cutis et al. (2011) [123]	VO	×	×	×	×	0.1	7.3	35K	26	–	–	×	✓	✓
Kim et al. (2013) [121]	VO	✓	×	×	×	–	–	1600	229.6	–	–	×	✓	✓
Golas et al. (2013) [127]	VO	×	✓	×	✓	0.01	3.5(5.5)[5]	1000	192.3	✓	–	×	×	✓
Kim et al. (2015) [122]	VO	✓	×	×	×	0.01-0.2	7.4	35K	5.7	–	–	×	✓	✓
Narang et al. (2015) [103]	VO	×	×	×	×	–	≈6.2	980	111.49	✓	✓	×	✓	✓
Narang et al. (2017) [184]	VO	×	×	×	×	–	2.5	200	185.2	–	✓	×	✓	✓
Sudkhot et al. (2023) [124]	VO	×	×	×	✓	0.25	–	40K	0.16	–	–	×	✓	✓
Burstedde et al. (2001) [12]	CA	×	×	×	×	0.3	(6.25)	–	–	–	–	×	✓	✓
Kirchner and Andreas (2002) [38]	CA	×	×	×	×	0.3	(6.25)	1116	–	–	–	×	✓	✓
Kirchner et al. (2003) [137]	CA	✓	×	×	×	0.3	(6.25)	–	–	–	–	×	✓	✓
Henein et al. (2004) [139]	CA	✓	×	×	×	0.3	(6.25)	–	–	–	–	×	✓	✓

Henein and White (2007) [140]	CA	✓	×	×	×	0.3	(6.25)	–	–	–	–	×	✓	✓
Guo et al. (2008) [130]	CA	✓	×	×	×	0.1	(6.25)	240	–	–	–	×	✓	✓
Ma et al. (2010) [185]	CA	×	×	✓	×	0.3	5.5(6.25)	–	–	✓	–	×	✓	✓
Zeng et al. (2011) [141]	CA	×	×	×	✓	0.3	(6.25)	200-300	–	–	–	×	✓	✓
Suma et al. (2012) [142]	CA	✓	×	×	✓	0.3	(6.25)	–	–	–	–	×	✓	✓
Bandini et al. (2014) [129]	CA	✓	×	×	×	0.33	7(12.5)	100-300	–	✓	–	×	✓	✓
Claudio et al. (2016) [134]	CA	✓	×	×	✓	–	6(10)	–	–	✓	✓	×	✓	✓
Lu et al. (2017) [138]	CA	×	✓	✓	×	0.3	(6.25)	750	–	–	✓	×	✓	✓
Jingwei et al. (2018) [128]	CA	×	×	×	×	–	8(8)	943	–	–	–	×	✓	✓
Fu et al. (2018) [131]	CA	✓	×	×	×	0.05-0.2	6(6.25)	1000	–	✓	✓	×	✓	✓
Zheng et al. (2019) [144]	CA	✓	×	×	×	1.0	(4.0)	300-800	–	–	–	×	✓	✓
Bazior et al. (2020) [132]	CA	✓	×	×	×	0.06-0.1	5.28(10.4)	–	–	–	–	×	✓	✓
Zhang et al. (2021) [186]	CA	×	×	×	✓	0.3	6.25(6.25)	–	–	✓	–	×	✓	✓
Huo et al. (2022) [135]	CA	✓	×	×	×	0.4	4.0(6.25)	180-600	–	–	–	×	✓	✓
Sarmady et al. (2022) [133]	CA	×	×	×	×	0.025	≈5.8(8)	–	–	✓	–	×	✓	✓
Zhang et al. (2023) [136]	CA	×	✓	×	×	0.33	1.77(6.25)	100	–	–	✓	×	✓	✓
Porzycki and Wąs (2023) [143]	CA	✓	×	✓	×	0.025 - 0.2	5	100-200	169.49	–	✓	×	✓	✓
Shao and Terzopoulos (2007) [64]	ABM	×	✓	✓	✓	–	–	1400	30	–	✓	✓	✓	✓
Robin et al. (2009) [149]	ABM	×	×	✓	×	–	–	–	–	–	✓	×	×	✓
Ondvrej et al. (2010) [150]	ABM	×	×	✓	×	–	–	200	25	–	–	×	×	✓
Samuel et al. (2012) [65]	ABM	×	×	✓	×	0.1	1.9	–	–	✓	✓	×	×	✓
Zhou et al. (2016) [147]	ABM	×	×	×	×	–	3	50-200	–	✓	–	✓	✓	✓
Godoy et al. (2016) [151]	ABM	×	×	✓	×	0.025	–	300	–	–	–	×	×	✓
Liang et al. (2016) [152]	ABM	×	✓	✓	×	–	–	100	161.3	–	–	×	×	✓

Liu et al. (2018) [153]	ABM	×	✓	✓	×	–	–	100-500	–	–	–	×	✓	✓
Mingbi et al. (2018) [19]	ABM	×	×	✓	×	–	–	–	–	✓	✓	×	×	✓
Sharma et al. (2018) [148]	ABM	×	✓	×	×	–	–	35-75	–	✓	–	✓	✓	✓
Li et al. (2021) [154]	ABM	✓	✓	×	✓	–	≈ 1.1	8-64	–	–	✓	×	✓	✓
Xie et al. (2022) [155]	ABM	✓	✓	✓	×	–	–	200	–	–	✓	×	×	✓
Alahi et al. (2016) [11]	Data-driven	×	×	×	×	0.4	–	–	–	–	✓	×	×	✓
Lee et al. (2017) [157]	Data-driven	×	×	×	×	0.4	–	–	–	–	✓	×	×	✓
Agrim et al. (2018) [66]	Data-driven	×	×	×	×	0.4	–	–	–	–	✓	×	×	✓
Anirudh et al. (2018) [158]	Data-driven	×	×	×	×	0.4	–	–	–	–	✓	×	×	✓
Amir et al. (2019) [159]	Data-driven	×	×	×	×	0.4	–	–	–	–	✓	×	×	✓
Zhu et al. (2022) [160]	Data-driven	×	×	×	×	0.4	–	–	–	–	✓	×	×	✓
Korbmacher et al. (2023) [9]	Data-driven	×	×	×	×	0.4	–	–	–	–	✓	×	×	✓
Dang et al. (2023) [7]	Data-driven	×	×	×	×	0.4	–	–	–	–	✓	×	×	✓
Hughes (2002) [14]	Macro	×	×	×	✓	–	5.0	–	–	–	–	×	✓	✓
Treuille et al. (2006) [15]	Macro	×	×	×	✓	–	–	10K	12	–	–	×	✓	✓
Narain et al. (2009) [67]	Macro	×	×	×	×	–	–	25K	11.35	–	–	×	✓	✓
Jiang et al. (2010) [163]	Macro	×	×	×	✓	–	5.0(10)	–	–	–	–	×	✓	✓
Jiang et al. (2015) [164]	Macro	×	×	×	✓	–	1.8(10)	–	–	–	–	×	✓	✓
Zhao et al. (2019) [173]	Macro	×	×	×	×	–	> 7	2000	–	–	–	×	×	✓
Jebrane et al. (2019) [170]	Macro	✓	×	×	×	0.01	7(8.5)	–	–	–	–	×	×	✓
van Toll et al. (2020) [166]	Macro	✓	×	×	×	0.02	> 8.0	30K	51.9	–	–	×	✓	✓
Yuan et al. (2020) [167]	Macro	×	×	×	×	0.0001	4	–	–	–	–	×	✓	✓
van Toll et al. (2021) [168]	Macro	✓	×	×	×	0.02	7.23(8)	30K	53.85	–	–	×	✓	✓
Liang et al. (2021) [174]	Macro	✓	×	×	✓	–	≈ 6.0(7)	–	–	✓	–	×	✓	✓
Chen et al. (2021) [171]	Macro	✓	×	×	×	–	–	2688	24.2	–	–	×	✓	✓

Zhou et al. (2022) [172]	Macro	×	×	×	×	–	–	–	–	–	–	×	✓	✓
Jiang et al. (2022) [187]	Macro	×	×	×	×	–	8(10)	3600	–	–	–	×	✓	✓
Li et al. (2023) [175]	Macro	×	×	×	×	–	4(7)	–	–	–	✓	×	✓	✓
Zhu et al. (2023) [176]	Macro	×	×	×	×	–	≈ 10	2000	–	–	✓	×	✓	✓
Xiong et al. (2009) [181]	Hybrid	×	×	✓	✓	–	≈ 4	4000	–	–	–	×	✓	✓
Xiong et al. (2010) [68]	Hybrid	×	×	×	×	–	≈ 2	–	–	–	–	×	×	✓
Anh et al. (2012) [69]	Hybrid	×	×	×	×	–	–	–	–	–	–	×	✓	✓
Xiong et al. (2013) [178]	Hybrid	×	×	×	×	0.5	≈ 6	1000	≈ 5.7	–	–	×	✓	✓
Rabiaa et al. (2016) [179]	Hybrid	×	×	×	×	–	–	–	–	–	–	×	✓	✓
Göttlich et al. (2018) [182]	Hybrid	✓	✓	×	×	0.0025	–	200	–	–	✓	×	×	✓
Yang et al. (2019) [70]	Hybrid	✓	✓	×	×	–	$\approx 5(5.4)$	–	–	–	–	×	×	✓
Biedermann et al. (2021) [180]	Hybrid	×	×	×	×	0.01	–	1000	–	–	–	×	×	✓
Dang et al. (2023) [3]	Hybrid	✓	✓	×	×	0.1	8	15000	≈ 1.4	–	–	×	✓	✓

Given the above comparison of operational-level model characteristics against various proposed criteria, assessments of the advantages and disadvantages of these models are summarized in Table 2.4 using the following rankings for each characteristic:

- $++$: The method can simulate the specified characteristic with minor or no changes.
- $+$: The method can simulate the specified characteristic but necessitates some modifications.
- $-$: The method can only simulate the specified characteristic through significant adaptations.
- $--$: The method cannot simulate the specified characteristic without substantial redesign.

[4] 2 fps on 6 cores.

[5] Max simulated and theoretical density: 3.5 and 5.5 $\mathrm{ped}/\mathrm{m}^2$.

Table 2.4: Assessments of the advantages and disadvantages of operational-level models for simulating high-density crowds (PI: physical interaction, GB: group behavior, FB: following behavior, RB: replan behavior, TSA: time step adaptivity, D: density, SS: simulation speed, S: scalability, IHM: integration of high-level modeling, AHM: ability of high-level modeling).

Models	**Behavior Capability**				**Model performance**				**Multi-level**	
	PI	GB	FB	RB	TSA	D	SS	S	IHM	AHM
SFMs	$++$	$++$	$++$	$+$	$-$	$+$	$+$	$-$	$--$	$++$
Advantages	Easy to integrate additional forces representing new pedestrian behavior into Newton's second equation 2.1				Stable with simulation time steps of 0.01 – 0.1s Simulating crowd densities of 2 – 6 $\mathrm{ped/m^2}$ Capable of simulating crowds at small scale (100 – 500 agents) and medium scale (500 – 5K agents)				Easy to be incorporated with strategic- and tactical-level models	
Disadvantages	Requiring careful calibration of parameters to avoid unrealistic behavior, particularly at extreme densities of 6 – 8 $\mathrm{ped/m^2}$				Sensitive to large simulation time steps of 0.2 – 0.4s High computational costs when simulating large-scale crowds					
VO models	$-$	$+$	$+$	$+$	$+$	$+$	$++$	$++$	$--$	$++$
Advantages	New pedestrian behavior can be addressed by establishing new velocity constraints and designing a new velocity obstacle space				Capable of adapting to a wide range of simulation time steps from 0.01 to 0.25s Simulating crowd densities of 2 – 6 $\mathrm{ped/m^2}$ Capable of simulating large-scale crowds (5K – 100K agents)				Easy to be incorporated with strategic- and tactical-level models	
Disadvantages					Stuck phenomena due to no collision-free velocities in extremely high-density conditions					
CA models	$-$	$-$	$-$	$-$	$--$	$++$	$-$	$-$	$++$	$-$
Advantages	New pedestrian behavior can be addressed by incorporating new floor fields				Capable of simulating crowd densities up to 8 $\mathrm{ped/m^2}$ Simulating crowds at small and medium scales					

Disadvantages	Addressing new pedestrian behavior increases the complexity of the model considerably	Simulation time step dependent on cell size High computational cost when simulating large-scale crowds	Strategic- and tactical-level models must be integrated in the form of floor field
ABMs	$+$ $+$ $+$ $++$	$+$ $+$ $-$ $-$	$-$ $++$
Advantages	Behavior-oriented modeling	Simulating crowds at small and medium scales	Easy to be incorporated with strategic- and tactical-level models
Disadvantages	Increasing complexity when simulating agents with complex decision-making capabilities	Simulation time step dependent on modeling technique High computational costs when simulating large-scale crowds	
Data-driven models	$--$ $--$ $--$ $--$	$--$ $--$ $-$ $-$	$--$ $-$
Advantages			
Disadvantages	Current lack of the availability of high-density pedestrian datasets	Short-term prediction of trajectory, typically 3 – 5s Prediction focusing only on one primary pedestrian Prediction time step dependent on the frame rate of training data	Cannot incorporate strategic- and tactical-level models
Macroscopic models	$-$ $-$ $-$ $+$	$-$ $++$ $++$ $++$	$++$ $--$
Advantages		Simulating smooth pedestrian movements due to small simulation time step Capable of simulating large-scale crowds	
Disadvantages	Focusing only on macroscopic crowd characteristics and phenomena	Sensitive to large simulation time steps	Strategic- and tactical-level models must be incorporated in the form of potential field

Hybrid models	+ + + +	− + + +	+ +
Advantages	Capacity to model new pedestrian behavior dependent on microscopic operational-level models integrated into the hybrid framework		Ability to incorporate strategic- and tactical-level models dependent on operational-level models integrated into the hybrid framework
Disadvantages		Synchronization of simulation time step between submodels	

In conclusion, certain operational-level models are commonly selected to address specific pedestrian behaviors and crowd phenomena in high-density scenarios. SFMs are widely used to model emergent behaviors of pedestrians in high-density situations and are capable of simulating medium-high density ranging from 2 to 6 $\mathrm{ped}/\mathrm{m}^2$. Upon proper calibration, these models can realistically simulate crowd density up to 8 $\mathrm{ped}/\mathrm{m}^2$. Next, CA models demonstrate the ability to simulate crowd densities from 2 to 8 $\mathrm{ped}/\mathrm{m}^2$, with extreme densities requiring finer discretization of the environment, whereas ABMs are primarily built to represent decision-making behavior such as following behavior. These three types of models are commonly used to simulate small- and medium-scale simulations with 100 to 5K agents. On the other hand, VO models demonstrate adaptability across different simulation time steps from 0.01 to 0.25s and are capable of simulating large-scale simulations with 5K - 100K agents. Data-driven models have limited capacity to simulate pedestrian behaviors in high-density situations, while macroscopic models are normally applied to crowded large-scale simulations. Finally, hybrid models are frequently employed for scenarios needing to investigate both crowd-level dynamics and individual-level interactions.

2.5 Future directions

Based on the previous evaluations of modeling approaches for pedestrian behaviors in crowded situations at different levels, this section proposes and discusses several potential research directions for the future development of dense crowd simulation. They include high-density data collection, deep learning perspectives, integration of different modeling levels, and multi-scale simulation.

High-density data collection

Most available datasets used for calibration and validation predominantly represent low-density situations. The ETH [188] and UCY [189] datasets, although widely used in this field, focus on low-density scenarios and do not adequately capture high-density situations. Conversely, only a limited number of high-density datasets have been collected through controlled experiments in Jülich [190], yet their laboratory setting imposes inherent limitations on their generalizability to real-world conditions. Thus, there is a need to collect more high-density datasets, particularly from real-world events to improve training and calibration of the models.

Deep learning perspectives

Data-driven models have received considerable attention and shown significant progresses in predicting pedestrian trajectories [28]. However, their deployment in high-density contexts remains a challenge. A major issue is the lack of high-density crowd datasets for training and testing neural network models under crowded situations. One promising solution to tackle this issue is using pedestrian simulation models to generate simulated data in high-density conditions, which can augment real datasets in training neural networks. Using both simulated data and real datasets together can increase the accuracy and generalizability of neural network models in various scenarios. Another promising direction is enhancing neural networks' predictive capacity by designing new neural network architectures or more effective training functions suited for high-density contexts. Additionally, novel evaluation metrics designed specifically for high-density scenarios must be proposed, together with classical distance-based accuracy metrics such as ADE and FDE, to accurately evaluate the performance of data-driven models in these situations.

Integration of different modeling levels

In complex environments with many obstacles, implementing only operational-level models may cause pedestrians to often get trapped in certain local areas. Thus, integrating multi-level modeling is required for such environments. One notable study integrates different techniques for modeling pedestrian behaviors at different levels into one framework, called the Menge framework [72]. However, despite its many strengths, this framework allows only one single modeling algorithm to be executed at each level during simulation.

In addition, most simulation models are developed to mimic one single behavior of pedestrians, which constrains the ability to represent the full range of pedestrian behaviors across various scenarios. Furthermore, pedestrian behavior also differs considerably between

low-density and high-density situations. Consequently, models that are effective and/or accurate for low-density environments may not effectively and/or accurately represent pedestrian behavior in crowded scenarios, and vice versa. Addressing this issue necessitates a general framework that integrates various models for pedestrian behaviors at different decisional levels. Such a framework needs the ability to switch dynamically between models and suitable transitioning rules for consistent and accurate simulations.

Multi-scale simulation

In large-scale real-world scenarios, it is common for participants to travel in social groups, often with friends and family members. The proportion of pedestrians moving in groups can vary between 40% and 70% [35, 34]. Furthermore, pedestrian flows tend to form naturally in densely populated areas where a large number of individuals share similar moving directions and velocities. Consequently, crowd simulation in these environments requires multi-scale entity representations, including the microscopic level for individuals, the mesoscopic level for groups, and the macroscopic level for flows, to manage the balance between detailed resolution and computational efficiency. Finally, the choice of simulation models, as well as when and where to apply these models, should be defined based on environmental information and crowd characteristics.

In addition, multi-scale simulations necessitate robust mechanisms for aggregation and disaggregation of entities to have a seamless transition between different scales. Entities in crowds, including individuals, groups, and flows, can merge or separate, as illustrated in Figure 2.12. Criteria to aggregate individuals into a group or a flow can be determined based on similarities in locations and velocities. Conversely, individuals may disaggregate from a group or a flow due to differences in targets, velocities, and objectives. Lastly, the formation of collective entities like groups and flows must be dynamically identified to apply appropriate models.

2.6 Summary and discussion

This chapter has given a comprehensive review of various approaches for modeling pedestrian behaviors with a primary focus on high-density scenarios. The systematic review investigated a total of 116 relevant articles published from 2000 to 2023. Different modeling techniques have been categorized into three classes corresponding to three decision-making levels: strategic, tactical, and operational. The analysis of collected articles has indicated that most studies focus

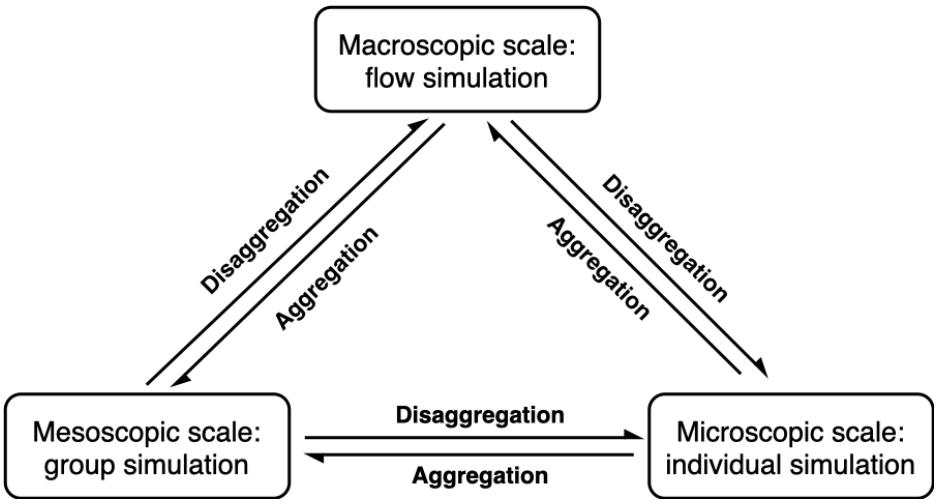


Figure 2.12: Multi-scale simulation.

on the operational level, which simulates the local pedestrian movement and interactions with other neighboring pedestrians and environments. These methods have been thoroughly examined, evaluated, and compared based on various criteria appropriate to high-density situations, such as model performance in simulating high-density levels, computational efficiency, extensibility, and ability to simulate emergent behavior in crowded conditions. Moreover, the advantages and disadvantages of these methods for high-density crowd simulation have been highlighted, providing researchers and modelers with critical insights to consider appropriate models for their study.

In addition, the systematic review has identified several existing research gaps in dense crowd simulations:

- The lack of high-density crowd datasets, especially in real-world conditions, currently restricts the accurate calibration and validation needed to enhance the realism of crowd simulations in high-density situations.
- Neural network data-driven approaches for predicting pedestrian trajectories have recently emerged in the last decade and have demonstrated accurate predictions. However, they mainly focus on low-density scenarios, characterized by their training functions that primarily address only distance-based loss, along with the lack of high-density crowd datasets.
- Specific models are frequently used to simulate particular aspects of crowd dynamics. However, there is a need for a comprehensive framework that combines these models

to cover a broader range of crowd characteristics across multiple scales, especially to address the dynamic variation of pedestrian behavior and movement in different density conditions.

Chapter 3

Data Collection

Contents

3.1 Festival of Lights 58
3.2 Experimental design 60
3.3 Methodology 61
3.3.1 Trajectory extraction method 61
3.3.2 Outflow calculation method 65
3.4 Dataset 66
3.4.1 Trajectory data 66
3.4.2 Outflow data 68
3.5 Validation 69
3.6 Data availability 70
3.7 Conclusion 71

The previous chapter reviewed state-of-the-art techniques for modeling pedestrian behaviors, with a primary focus on high-density situations. It also highlighted existing research gaps for these situations, notably the need for more high-density datasets of pedestrian movements in real-world conditions. This chapter aims to address this gap by introducing a new high-density dataset collected during a mass-gathering event. The experiment to collect this dataset was conducted under real-world conditions together with members of the MADRAS project, and the associated research paper is currently in submission [8].

The chapter is organized as follows. Section 3.1 introduces an overview of the history of the Festival of Lights. Section 3.2 describes the experimental setup of various cameras to capture crowd movements during the festival. Section 3.3 presents tools and methods for extracting trajectory and pedestrian outflow data from numerous recorded videos. Section 3.4 showcases a dataset collected during the festival, along with basic statistics and analyses of this dataset. Next, Section 3.5 discusses the validation of the collected data, and Section 3.6 provides information about its availability for public use. Finally, the chapter concludes in Section 3.7.

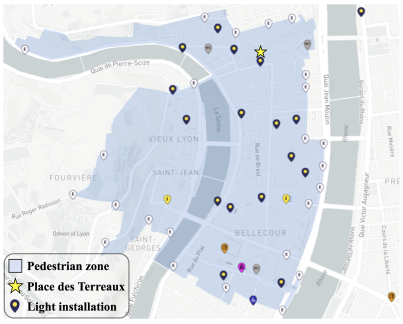
3.1 Festival of Lights

The Festival of Lights [191], also known as la Fête des Lumières in French, is a spectacular annual festival held in Lyon, France. It takes place around the 8^{th} of December and lasts over four nights. The festival was originally rooted in tradition and faith when the inhabitants of Lyon erected a statue of the Virgin Mary for protection against recurrent floods and social unrest. Although the inauguration was initially scheduled for September 8, 1852, flooding caused it to be postponed to December 8 of that year. On the rescheduled date, the citizens illuminated their windows and balconies with candles to display unity and solidarity. This tradition continues today and has become an international festival attracting millions of visitors yearly.

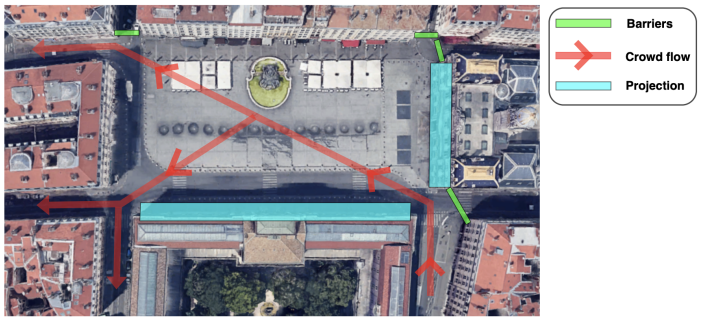
Nowadays, the Festival of Lights is not merely a tribute to the Virgin Mary but also a vibrant showcase of light arts. During the festival, the city is transformed into an open-air gallery of light arts with animations projected onto the facades of ancient buildings and historical monuments. The festival typically spans four nights (from December 8^{th} to 11^{th} in 2022), with light displays active usually from 7 PM to midnight.

In 2022, the Festival of Lights attracted approximately two million participants over a total of 30 artworks displayed across the festival area (as shown in Figure 3.1a) in the metropolis of Lyon. The focal point of attraction was Place des Terreaux – the city's central plaza with an area of around 7000 square meters. As the festival's largest and most crowded place, it welcomed about 150000 people each night. Figure 3.1b describes the main circulation of crowd

movements in Place des Terreaux. Crowd flows are pre-configured by the festival organizer to be unidirectional on a macroscopic scale to ensure safety and prevent congestion. Pedestrians are only permitted to enter the plaza through a single entrance (President Edouard Herriot road) from the southeast. Exiting the plaza is possible via three roads: Algeria road from the northwest, and Chenavard road and Constantine road as the primary outflow from the southwest. The remaining roads adjacent to the plaza are blocked by barriers to prevent entry or exit and maintain the intended unidirectional flow. Gatekeepers are assigned at the entrance to monitor and control the number of people entering the plaza, ensuring it does not exceed approximately two-thirds of the plaza's capacity by closing a barrier to restrict further access before the start of each light show. Figure 3.1c shows a light show in Place des Terreaux with thousands of people gathering to watch the show.



(a) Festival area[6].



(b) Main circulation of the crowd flow at Place des Terreaux.

(c) A screenshot of pedestrians gathering at Place des Terreaux to watch a light show[7].

Figure 3.1: Festival of Lights in 2022.

[6] `https://www.fetedeslumieres.lyon.fr/en/map/2022-map` (accessed on 31 May 2024).

[7] `https://www.fetedeslumieres.lyon.fr/fr/oeuvre/grand-mix-au-musee-des-beaux-arts-de-lyon` (accessed on 31 May 2024).

Given that Place des Terreaux is the central and most crowded location of the Festival of Lights, our data collection experiment focuses particularly on this area. The description of the experiment is presented in the following sections.

3.2 Experimental design

The objective of our experiment is to capture the dynamics of dense crowds during the Festival of Lights from various perspectives. Therefore, different cameras are installed to record crowd movement at both microscopic and macroscopic scales. Figure 3.2 illustrates the experimental design at Place des Terreaux, detailed as follows:

- Camera 1 (CCTV camera) is set atop the city hall (east side) to provide the overall large-scale view of crowd movement in Place des Terreaux.

- Cameras 2 and 3 (SJCAM A10 cameras) on the northeast side are mounted on the balcony of an apartment on the second floor to capture the top-down views of the street below. The purpose of these cameras is to focus on a small area with high resolution for tracking pedestrian trajectories from the videos.

- Cameras 4 and 5 (SJCAM A10 cameras) are set to record the outflows on Chenavard road and Constantine road, respectively.

- All cameras are configured to operate in night vision mode to ensure video quality during nighttime recording sessions.

Figure 3.3 shows screenshots of all cameras. Approximately 200 GB of videos were recorded from the 8^{th} to the 10^{th} of December 2022. In addition to video data, other information types are collected, including collision counts, GPS data, and surveys. However, these additional data are out of the scope of this study and will not be discussed in the manuscript. The primary focus of this study is to present two specific types of pedestrian data extracted from video footage: pedestrian trajectories and pedestrian outflow. The pedestrian outflow data is used to calibrate and validate a hybrid framework for crowd simulation developed in Chapter 4, whereas the trajectory dataset is used to train a neural network developed in Chapter 5.

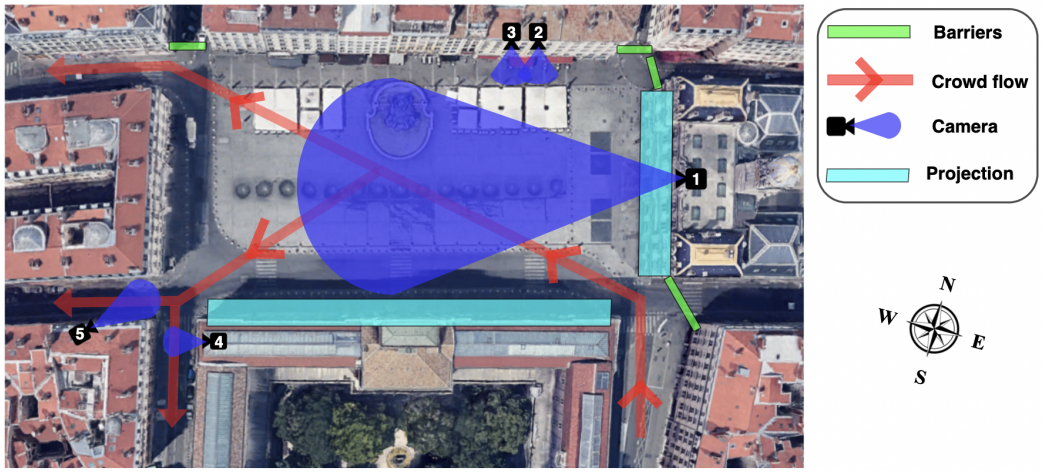


Figure 3.2: Experimental design.

3.3 Methodology

3.3.1 Trajectory extraction method

This section presents the process for extracting pedestrian trajectories from video footage, including the choice of an appropriate tool, camera calibration, and a step-by-step trajectory tracking procedure.

Tool

Compared to controlled experiments conducted under laboratory conditions [192], our experiment inevitably encounters the following difficulties:

- Manual measurements in real-world conditions may not be as precise as those in laboratory settings due to weather conditions.
- The experiment is conducted in an outdoor setting at night time, resulting in uneven lighting across the recording area. This inconsistency in the lighting sometimes made it difficult to distinguish pedestrian features, such as pedestrian heads, from the background.
- Sudden changes in illumination during the festival, such as transitions from dark to bright and vice versa, can cause false detections of pedestrians.
- High-density levels of crowds led to significant occlusion and overlapping of pedestrian shapes, causing pedestrians to block each other from view, making it challenging to

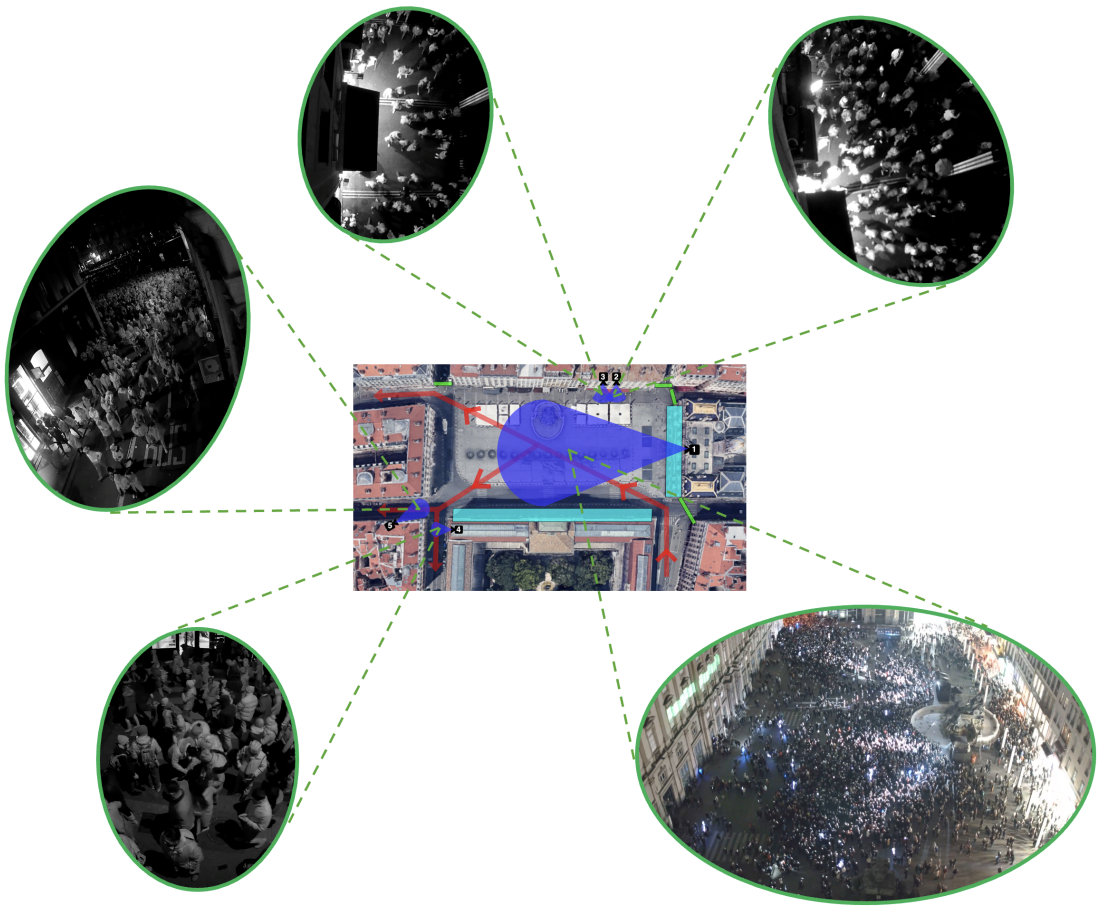


Figure 3.3: Screenshots of all cameras.

identify each person separately.

- Complex behaviors of pedestrians in outdoor and natural environments such as running pedestrians and temporarily standing people/groups.
- People with additional objects such as umbrellas, balloons, umbrellas, chairs for babies, etc.

Because of these aforementioned challenges, the performance of algorithms for fully automated pedestrian detection and tracking in this situation is significantly reduced. Therefore, we decided to use PeTrack [193, 194], a software that provides semi-automated tracking of pedestrian trajectories from video footage, to facilitate the task of extracting pedestrian trajectories. This software offers the following main functionalities:

- Camera calibration: PeTrack provides intrinsic and extrinsic camera calibration functions to map 2-dimensional image coordinates to real-world 3-dimensional coordinates.
- Pedestrian detection: PeTrack can automatically detect pedestrians using special markers (e.g., colorful hats that are easy to recognize). This function generally works best in controlled laboratory experiments with ideal settings, where participants are asked to wear such colorful markers. However, in real-world experiments, these conditions are usually not met. Therefore, manual detection methods for pedestrian detection are typically used in real-world conditions.
- Automatic tracking: PeTrack employs the pyramidal iterative Lucas-Kanade feature tracker [195] for automatically tracking detected pedestrians. This method remains effective even for experiments conducted in real-world outdoor settings where weather and lighting conditions are usually not ideal.
- Manual correction: Users can manually correct any errors in tracked trajectories to ensure data quality.

Details on the trajectory extraction process using PeTrack are presented in the next sections.

Camera calibration

Camera calibration is an essential preprocessing step in computer vision applications that estimates camera parameters to accurately map two-dimensional image coordinates to three-dimensional (3D) real-world coordinates. This process consists of two main tasks: intrinsic calibration and extrinsic calibration. Intrinsic calibration determines the internal parameters of the camera (focal length, geometric distortion, optical center, skew coefficient) by using images of a calibration pattern such as a chessboard. On the other hand, extrinsic calibration calculates the camera's position and orientation in the 3D scene. It determines external parameters (rotation matrix and translation vector) that transform points from the real-world 3D coordinates to the 3D camera coordinates.

The camera calibration for our data collection experiment is performed using PeTrack software, which provides an automatic tool for this task based on Zhang's method [196]. The camera's internal parameters are obtained by importing multiple images of a chessboard (with the chessboard covering most of the image area) captured from various angles into PeTrack software for automatic calculation. For extrinsic calibration, a human wearing a white hat served as the reference pattern. The human stands at six positions within the recording area,

forming a total of 11 reference points (5 points for the head and 6 midpoints between two feet), as illustrated in Figure 3.4. Once both intrinsic and extrinsic parameters are calculated, they are input into the tracking step.

Figure 3.4: Image from camera 3 including reference points for extrinsic calibration.

Trajectory tracking

Video footage from cameras 2 and 3 is used to extract pedestrian trajectories. Chunks of videos corresponding to different density levels (from medium to high) are selected for data extraction. The following steps outline the workflow to track pedestrian trajectories from our videos using PeTrack software:

- Define a region of interest for tracking pedestrian trajectories. An area of approximately 10m x 6m is designated by drawing yellow lines into the videos, as shown in Figure 3.5. The region of interest is selected to be consistent in videos recorded by both cameras 2 and 3. Only pedestrian trajectories appearing inside this region would be tracked.

- Manually detect by clicking pedestrian heads when they enter the region of interest.

- Automatic tracking is applied for each detected pedestrian using the pyramidal iterative Lucas-Kanade feature tracker [195] integrated into PeTrack.

Figure 3.5: Region of interest defined by a yellow rectangle for videos recorded by camera 3.

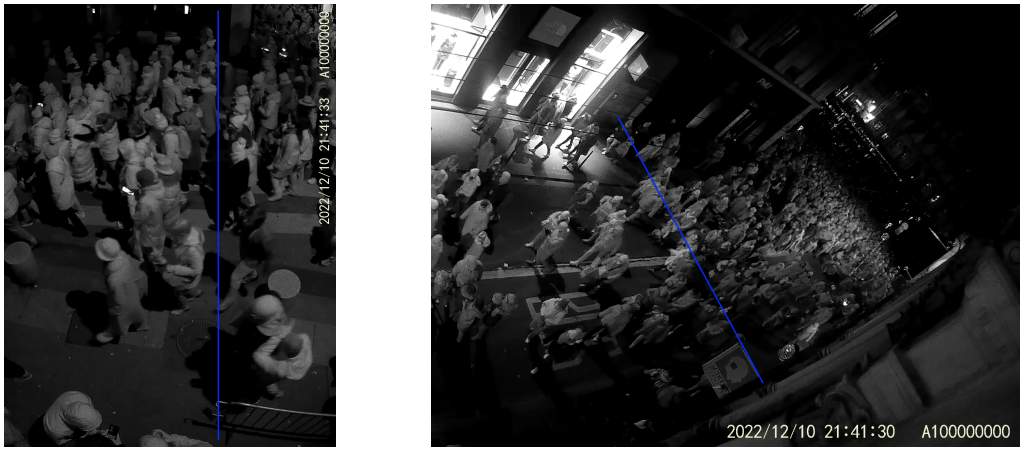
- Visually monitor the tracking of pedestrian trajectories and manually recalibrate the trackers to align with corresponding pedestrian heads whenever the trackers are lost or produce distance errors. This particularly happens when the illumination suddenly changes, or pedestrians transition between uneven lighting areas.
- Save trajectories as soon as pedestrian heads move out of the region of interest.

Once the tracking process is completed, PeTrack generates a .txt file as output, which can be re-imported back to PeTrack. This file begins with a header of comments providing relevant information such as fps, explanations of column variables, etc. Following the header is the data section where each row represents coordinate data for each pedestrian at each time frame.

3.3.2 Outflow calculation method

Given the aforementioned difficulties in data collection and extraction, the manual counting method is used to ensure accuracy. This method quantifies the pedestrian outflow on Chenavard and Constantine roads, corresponding to videos recorded by cameras 4 and 5, respectively. Due to the labor-intensive nature of this method, the calculation focuses only on a short recording time frame from 9:38 PM to 9:45 PM on December 12, 2022. This time frame corresponds to the exit of crowds from the plaza after the end of a show at Place des Terreaux. To ensure accurate counting, a reference line is drawn across each road in the recorded videos (as illustrated in

Figure 3.6), and the pedestrian count is incremented each time a pedestrian's head crosses this reference line. The outflow data is then used to simulate the crowd exit scenario at the Festival of Lights in Chapter 4.



(a) Chenavard road. **(b)** Constantine road.

Figure 3.6: Reference lines to count pedestrian heads.

3.4 Dataset

This section presents the trajectory and outflow data extracted from pedestrian recordings as well as the analysis of these data.

3.4.1 Trajectory data

Basic statistics

A total of 9 video segments recorded on December 8, 2022, are selected for extraction, with 3 of them recorded by camera 2 and 6 recorded by camera 3. Data corresponding to each video segment is labeled according to the camera and time (e.g., three data records from camera 2 are named TopView_1A, TopView_1B, and TopView_1C, where "TopView" denotes the top-view scene captured by cameras 2 and 3). Furthermore, these data are categorized into 3 classes based on the type of pedestrian movements observed:

- Unidirectional flow: TopView_1A, TopView_2A, and TopView_2B.
- Unidirectional flow with standing groups as obstacles: TopView_1B and TopView_2C.

- Bidirectional flow: TopView_1C, TopView_2D, TopView_2E, and TopView_2F.

Table 3.1 presents the basic statistics of the TopView dataset, including timestamps, trajectory counts, mean density, maximum density, mean speed, and mean trajectory duration. The data records cover the time range from 9:26 PM to 9:56 PM. A total of 5269 trajectories have been extracted from 9 video segments, with average densities ranging from 0.37 to 1.58 $\mathrm{ped}/\mathrm{m}^2$. Based on the above classification and the information from Table 3.1, it can be clearly seen that unidirectional flow scenarios generally achieve higher density levels compared to bidirectional flow scenarios. Additionally, Table 3.1 indicates clear correlations between mean density and both mean speed and mean trajectory duration, with higher density associated with lower speed and longer trajectory duration.

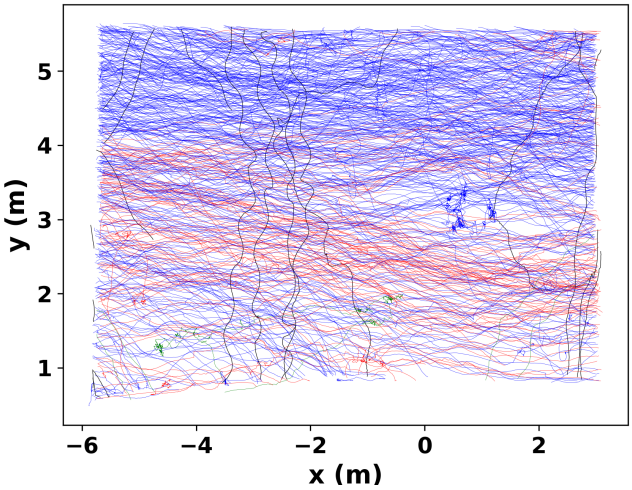
File	**Start** [UTC+1]	**End** [UTC+1]	**Traj**	**Mean density** [ped/m^2]	**Max density** [ped/m^2]	**Mean speed** [m/s]	**Mean trajectory duration** [s]
TopView_1A	22:40:45	22:44:15	965	1.13	1.54	0.57	12.27
TopView_1B	22:55:06	22:57:46	685	1.07	1.44	0.52	12.64
TopView_1C	23:10:33	23:13:58	673	0.65	1.07	0.78	9.03
TopView_2A	21:26:27	21:29:07	711	1.58	2.29	0.41	18.11
TopView_2B	21:40:39	21:43:25	612	1.11	2.58	0.58	13.18
TopView_2C	22:55:18	22:58:20	693	1.06	1.32	0.54	12.94
TopView_2D	23:10:16	23:12:57	529	0.64	1.04	0.75	8.96
TopView_2E	23:24:59	23:26:11	218	0.57	1.00	0.75	8.63
TopView_2F	23:54:59	23:56:30	183	0.37	0.65	0.95	6.92
		Total	5269	0.96	1.44	0.63	12.22

Table 3.1: File tracking, timestamps, trajectory counts, and main statistics for the TopView dataset.

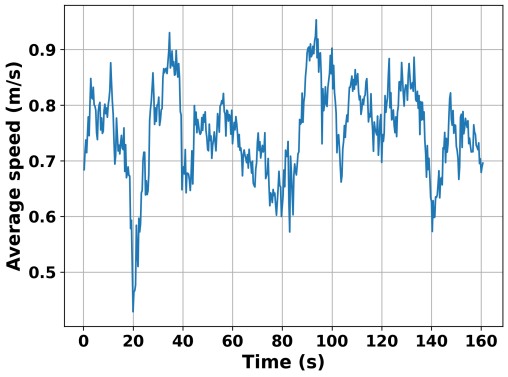
More detailed analysis of bidirectional flow data

This study particularly focuses on bidirectional flow data as they represent more interactions and collision avoidance behavior between pedestrians in crowded situations and will be used in the next chapters. Given this, the TopView_2D data is particularly selected for detailed analysis.

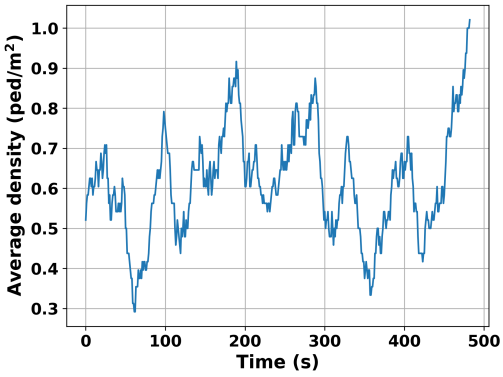
A total of 529 trajectories are collected for TopView_2D, with 396 moving to the left, 109 moving to the right, and 13 upward and 11 downward, as illustrated in blue, red, black, and green, respectively, in Figure 3.7a. The average speed ranges from 0.43 to 0.95 m/s and the average density fluctuates between 0.3 and 1.0 $\mathrm{ped}/\mathrm{m}^2$, as shown in Figure 3.7b and 3.7c.



(a) Trajectories in blue: moving left, red: moving right, black: moving upward, and green: moving downward.



(b) Average speed over time.



(c) Average density over time.

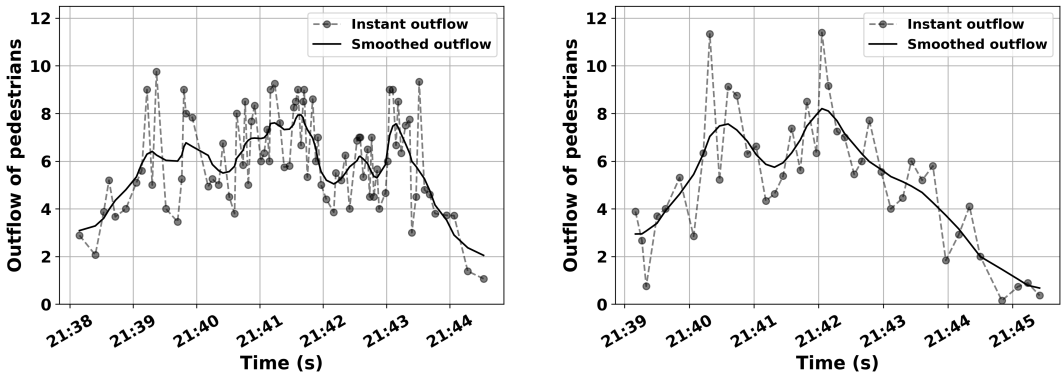
Figure 3.7: TopView_2D

3.4.2 Outflow data

There are 3833 pedestrians counted in the videos, with 2030 pedestrians on Paul Chenavard road and 1803 pedestrians on Constantine road, respectively. The instant pedestrian outflow is computed by dividing the number of pedestrians crossing the reference line in a given duration by the corresponding time interval. The black dots in Figure 3.8 represent the instant outflows, which peak at nearly 10 ped/s on Chenavard road and over 11 ped/s on Constantine road. Crowd flows last approximately 6.5 minutes until very few passes of pedestrians are counted. However,

not all pedestrians leave immediately after a projection show ends, as some remain in the plaza to watch subsequent shows.

Additionally, to improve the accuracy of the data and mitigate noises, the instant outflow is smoothed using the Gaussian filter method [197] with a standard deviation of 2.0 for the kernel and is presented by the solid black lines in Figure 3.8.



(a) Chenavard road. **(b)** Constantine road.

Figure 3.8: Time-series outflow of pedestrians on Chenavard road and Constantine road.

3.5 Validation

The extracted trajectory dataset is validated through the manual cross-checking process, with independent verification (visual inspection and corrections when necessary) by a project member who is not involved in tracking the corresponding data. Furthermore, given that the regions of interest for extracting pedestrian trajectories are designated to be similar in videos recorded by cameras 2 and 3, and time is synchronized across all cameras, two series of the same scene are selected for further comparison. Specifically, two pairs of videos, including TopView_1B versus TopView_2C and TopView_1C versus TopView_2D, are chosen as they have partially overlapping timestamps (as indicated in Table 3.1). Figure 3.9 illustrates an example of the same scene in TopView_1B and TopView_2C videos, both at 22:55:25.

An additional step for validating the measurement is implemented by comparing the similarity of trajectory data extracted from these two pairs of videos. Figure 3.10 presents the superposition of average speed and density sequences over time. The root mean square error (RMSE) is used to evaluate the differences between the mean speed and density time series in the overlapping sequences. The RMSE results, as shown in Table 3.2, indicate relatively small differences,

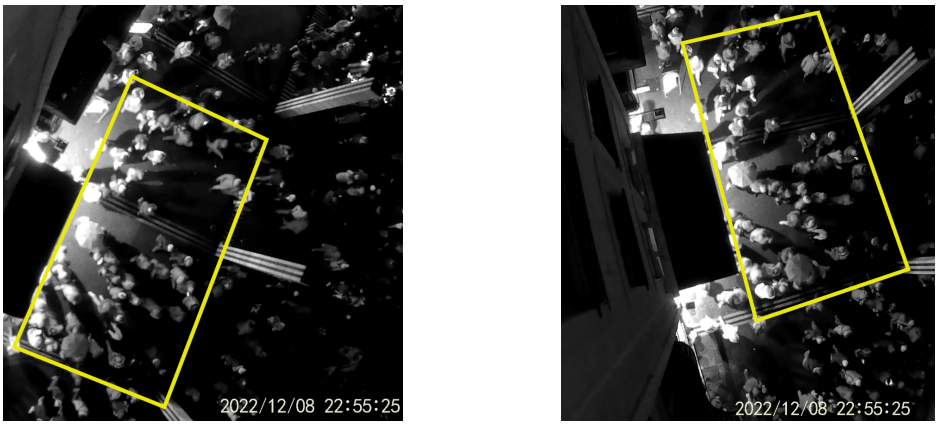


Figure 3.9: Similar regions of interest in the TopView_1B and TopView_2C camera videos, which capture the same scene but from different locations and angles.

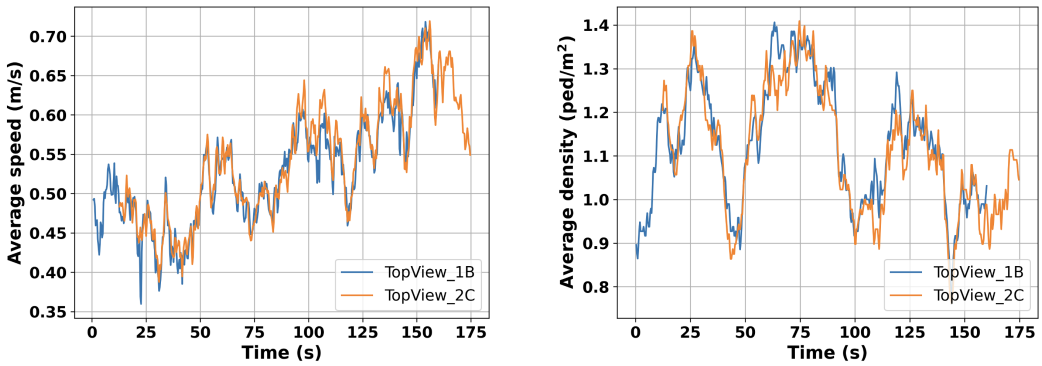
ranging from 5 and 10% of the estimated average speed and density. A systematic error occurs in the measurements but in opposite directions for the TopView_1B/TopView_2C recordings and the TopView_1C/TopView_2D recordings. Specifically, the average density over time of TopView 1B is slightly higher than that of TopView 2C, while the average density of TopView 1C is lower than that of TopView 2D.

	RMSE mean speed $[\mathrm{m/s}]$	RMSE density $[\mathrm{ped/m^2}]$
TopView_1B / TopView_2C	0.02	0.05
TopView_1C / TopView_2D	0.05	0.07

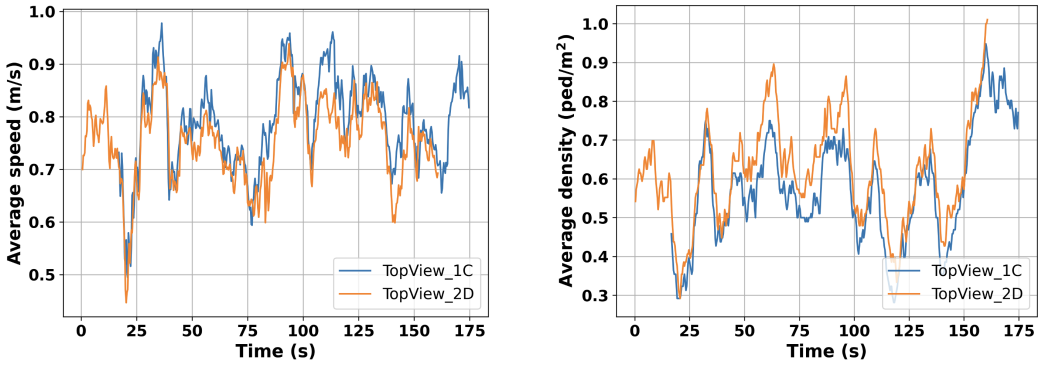
Table 3.2: Root mean square errors (RMSE) between the pedestrian mean speed and density time series for the TopView_1B/TopView_2C and TopView_1C/TopView_2D video recordings.

3.6 Data availability

The data is publicly available on `https://madras-data-app.streamlit.app` - an online platform developed for providing the visualization, analysis, and interactive exploration of the datasets. The source code of this application is hosted on GitHub [198] and is distributed under the MIT license.



(a) TopView_1B/TopView_2C.



(b) TopView_1C/TopView_2D.

Figure 3.10: Partial overlap of the average speed and density sequences over time for two pairs of video recordings: TopView_1B/TopView_2C and TopView_1C/TopView_2D, which capture the same scene from different positions and perspectives.

3.7 Conclusion

This chapter presents the experiment conducted at the Festival of Lights 2022 to collect high-density pedestrian data. This data collection was conducted in collaboration with members of the Franco-German MADRAS project team. The author made the following contributions: participating in the experimental setup, calibrating cameras, tracking two trajectory sub-data: TopView_2A and TopView_2E, extracting pedestrian outflow, and writing a part of the article manuscript. The data covers both macroscopic and microscopic scales, including outflow and trajectories of pedestrians. This contribution addresses the current lack of publicly available data in high-density real-world scenarios, as most existing dense crowd data originates from controlled laboratory experiments. The dataset is released publicly online through an online

application. Subsequently, this dataset is used to improve two types of models: simulation models and prediction models, which are described in the next chapters.

Chapter 4

HyPedSim – A Hybrid Pedestrian Simulation Framework

Contents

4.1 Introduction 74
4.2 HyPedSim Framework 76
4.2.1 General Overview 76
4.2.2 Agent-Based Model for Multi-Level Behaviour 77
4.2.3 Pedestrian Activity Diagram 80
4.3 Application to the Festival of Lights 80
4.3.1 Simulation scenario 82
4.3.2 Pedestrian Simulation Models 83
4.3.3 Simulation Details 88
4.3.4 Model Calibration 88
4.3.5 Calibration Results 92
4.3.6 Model Validation 94
4.3.7 Sensitivity Analysis 94
4.4 Performance Analysis 98
4.5 Conclusions and Discussion 101

The previous chapter has introduced the Festival of Lights, an annual mass gathering event, and has presented the data collected during this event. This dataset addresses the lack of high-density pedestrian data in real-world conditions. In this chapter, we aim to investigate the use of density-related factors in hybrid approaches to develop a comprehensive pedestrian simulation framework capable of simulating various crowd scenarios. Moreover, we explore how the high-density empirical data presented in Chapter 3 is used to calibrate this framework. The work in this chapter has been presented in PED23 [1] and PAAMS 2023 conferences and published in [3, 4] where the author was the primary contributor.

4.1 Introduction

Pedestrian modeling and simulation is an active research field that models and predicts how pedestrians move and behave in different environments. Generally, pedestrian simulation models are categorized into three levels [29]: strategic-level models (also known as goal selection models) [73, 78], tactical-level models (or path planning models) [56, 59], and operational-level models [13, 63]. Definitions and detailed reviews of these models were presented in Chapter 1 and Chapter 2. However, studies in the literature primarily focus on the operational level because this level refers to immediate, physical actions and reactions of pedestrians, which are the most fundamental aspects in modeling pedestrian behaviors [6]. This study aims to investigate the use of density-related factors to develop a comprehensive simulation framework for pedestrian behavior, with a primary focus also on the operational level.

Various operational-level models have been developed in the literature, each designed to simulate and mimic specific phenomena [46]. Microscopic models, which focus on simulating individual behavior and interactions, are commonly used in low- to medium-density situations. These models have been known to exhibit specific limitations in high-density scenarios and thus require careful calibration to be applicable in such situations. For example, several drawbacks of microscopic models in high-density scenarios have been reported, such as abnormal oscillation in trajectories with the Social Force Model [18], congestion in dense bidirectional flow with the Velocity Obstacle Model [19], and predicted trajectories with many collisions with the data-driven models [7, 9]. Conversely, macroscopic models [199, 15], which simulate crowds as a whole, are typically used for high-density scenarios due to the assumption that pedestrians in crowded scenarios tend to move in a continuous and collective manner.

In order to provide a generic framework able to model a wide range of situations in terms of density, it thus appears necessary to combine several approaches. I thus follow and propose a hybrid modeling approach. Hybrid modeling typically combines macroscopic or mesoscopic

models with microscopic models into one framework to exploit the strengths of both approaches. This achieves a balance between simulating the overall crowd movement and behaviors with reasonable computational cost while maintaining detailed representations of individual behaviors and interactions. A review of related works on hybrid modeling approaches for pedestrian simulation was presented in Chapter 2. Although Curtis et al. [72] proposed the Menge framework, a crowd simulator that aggregates different modeling algorithms at each pedestrian behavioral level, it can only execute one modeling technique at the operational level during a simulation. Currently, a notable research gap is the lack of a comprehensive framework capable of dynamically switching between models to adapt to various crowd situations.

This chapter proposes and evaluates an agent-based framework for pedestrian simulation that allows agents to dynamically switch their operational-level models in response to changes in local density states to adapt to different crowd scenarios in the environment, hence the name HyPedSim framework. The agent-based approach is chosen to design the architecture of the framework due to the complex architecture that captures various models with the flexibility to switch. The ability of the framework to switch between operational-level models is illustrated through an application of dense crowd simulation at the Festival of Lights in Lyon, France. Moreover, the proposed framework also demonstrates that dynamic switching can be extended to higher levels of modeling pedestrian behaviors, such as strategic and tactical levels. The simulation environment is divided into predetermined areas, with the local density of each area estimated based on expert knowledge rather than explicit computation, making it more intuitive and easier to implement. Based on the descriptions of hybrid models in Chapter 2, our framework can be classified as a region-based model. Compared to region-based models presented in Chapter 2, our framework introduces the following novel features:

- Flexibility to create various combinations of models for different areas of the environment.
- Extensibility of dynamic switching among models to higher-level modeling of pedestrian behaviors, such as the strategic and tactical levels.

The chapter is structured as follows: Section 4.2 proposes a framework that provides the ability to capture various models at different behavioral levels of pedestrians and allows for dynamic switching of these models. The framework's ability to couple different models is presented in Section 4.3 through an application simulating pedestrian dynamics at the Festival of Lights in Lyon, France. Section 4.4 evaluates and compares the performance of the coupled model. Finally, Section 4.5 provides a conclusion and a discussion of the study.

4.2 HyPedSim Framework

This section proposes HyPedSim, a framework for simulating pedestrian dynamics that provides the ability to integrate various models and allows for dynamic switching between them to simulate different crowd situations. It begins with a general overview of the framework, followed by detailed descriptions of the architecture design and the agent's activity diagram. What should be noted is that while our primary focus is on the operational level, the tactical level is also incorporated into the framework's architecture to demonstrate its extensibility in modeling the higher level of pedestrian behavior.

4.2.1 General Overview

The environment is discretized into separate zones, each containing walkable space for pedestrians and potential obstacles. This partitioning can be informed by expert knowledge or based on specific conditions such as crowd characteristics and environmental data. Figure 4.1 illustrates an example of the environment divided into four different zones, while Section 4.3 presents a concrete example of environment partitioning for a case study. Pedestrian crowds in each zone are simulated by an appropriate simulation model, which can be selected depending on the crowd density in this zone.

Each zone is associated with a specific transition function. The transition function, activated when pedestrians enter or exit the zone, aggregates or disaggregates the information of these pedestrians to ensure a smooth transition between zones.

Pedestrians are modeled at the individual level. They can move from one zone to another. At each behavioral level, each pedestrian is simulated using a specific model that depends on the zone they belong to. When pedestrians enter a new zone, they dynamically update their attributes to align with models used in the new zone.

The next section proposes a generic and extensible method for this update process. Specifically, an agent-based architecture is designed to capture various models and enable pedestrian agents to dynamically change their operational-level model based on zone-specific information. While the agent-based architecture is primarily designed for dynamic switching between operational-level models, the strategic and tactical levels are also integrated to demonstrate its extensibility to higher levels of modeling pedestrian behaviors.

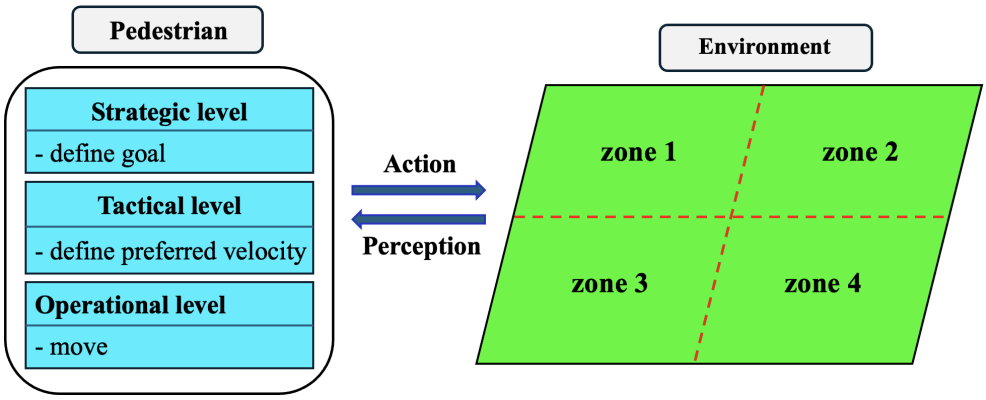


Figure 4.1: General overview.

4.2.2 Agent-Based Model for Multi-Level Behaviour

Firstly, the problem is formulated as a mathematical abstraction and divided into three sub-problems: *strategic*, *tactical*, and *operational subproblem*. Similar formulations for crowd simulation problems can be found in previous works [72].

Let $\mathbb{S} = \{E, A\}$ represent the simulation state, including the environment E and the set of agents A. The *strategic subproblem* is formulated as follows:

$$S : \mathbb{S} \times t \rightarrow \mathbb{R}^2 \tag{4.1}$$

For each agent, the function S maps the simulation state $\mathbb{S}$ and time t to a point representing a goal in $\mathbb{R}^2$.

The *tactical subproblem* can be defined as follows:

$$T : \mathbb{S} \times t \times \mathbb{R}^2 \rightarrow \mathbb{R}^2 \tag{4.2}$$

The function T maps the simulation state, time, and the agent's goal to a preferred velocity in $\mathbb{R}^2$. This preferred velocity represents the direction extracted from either a global or local path that pedestrians use to move toward their goal.

The *operational subproblem* can be represented as follows:

$$O : \mathbb{S} \times t \times \mathbb{R}^2 \rightarrow \mathbb{R}^2. \tag{4.3}$$

The function O maps the simulation state S, time, and the agent's preferred velocity to a feasible velocity in $\mathbb{R}^2$. The feasible velocity is then used as the agent's next velocity and applied to

update the agent's position in the next simulation step.

The three *subproblems* can be generally incorporated into the following mathematical formulation:

$$\mathbf{v}_i(t) = O_i(\mathbb{S}, t, T_i(\mathbb{S}, t, S_i(\mathbb{S}, t))) \tag{4.4}$$

$$\mathbf{p}_i(t+1) = \mathbf{p}_i(t) + \mathbf{v}_i(t)\Delta t \tag{4.5}$$

where $\mathbf{v}_i(t)$ and $\mathbf{p}_i(t)$ represent the velocity and the position of agent i in the xy-coordinate system at time t, respectively, while Δt is the duration of a simulation step.

Based on the above mathematical formulation, an agent-based model is designed to capture various models and provide agents with a generic way to switch between these models. Figure 4.2 presents the architecture of the agent-based model. In the architecture, pedestrian agents are represented by the ***Pedestrian*** class. Note that "class" in this context refers to the classical definition in the Object Oriented Programming paradigm. The ***Pedestrian*** class has three key attributes, $\mathrm{strategic\_level}$, $\mathrm{tactical\_level}$, and $\mathrm{operational\_level}$, along with and three main methods: *define goal*, *define preferred velocity*, and *move*.

Furthermore, each *subproblem* is represented by an interface, hence the names ***Strategic level, Tactical level***, and ***Operational level*** for these interfaces. Each interface is responsible for capturing different inheriting subclasses, with each subclass corresponding to a specific model that simulates pedestrian behavior at the same level as the interface (strategy design pattern [200]). The ***Strategic level*** and ***Tactical level*** include abstract methods for computing a goal and a preferred velocity for each pedestrian, respectively, whereas the ***Operational level*** interface has an abstract method for calculating their next velocity. Any model inheriting from an interface must implement the methods specified in that interface. For example, the Social Force Model [13] can be a subclass inheriting from the ***Operational level*** interface, and thus it must implement a method for calculating the next velocity for pedestrian movement in the next simulation step.

The three methods in the ***Pedestrian*** class, *define goal*, *define preferred velocity*, and *move*, use the three attributes $\mathrm{strategic\_level}$, $\mathrm{tactical\_level}$, and $\mathrm{operational\_level}$, respectively, to communicate with subclasses of ***Strategic level***, ***Tactical level***, and ***Operational level*** interfaces. Pedestrian agents can only use one model at any given time for each behavioral level; however, they can switch models based on environmental conditions and crowd characteristics. Switching between models is made by simply updating their corresponding attributes. For example, if an agent wants to switch their operational-level model from "model 1" to "model N", it sets the $\mathrm{operational\_level}$ attribute to the new value "model N". Consequently, when the method *move*

invokes the operational_level attribute with the updated value, the "model N" is then executed.

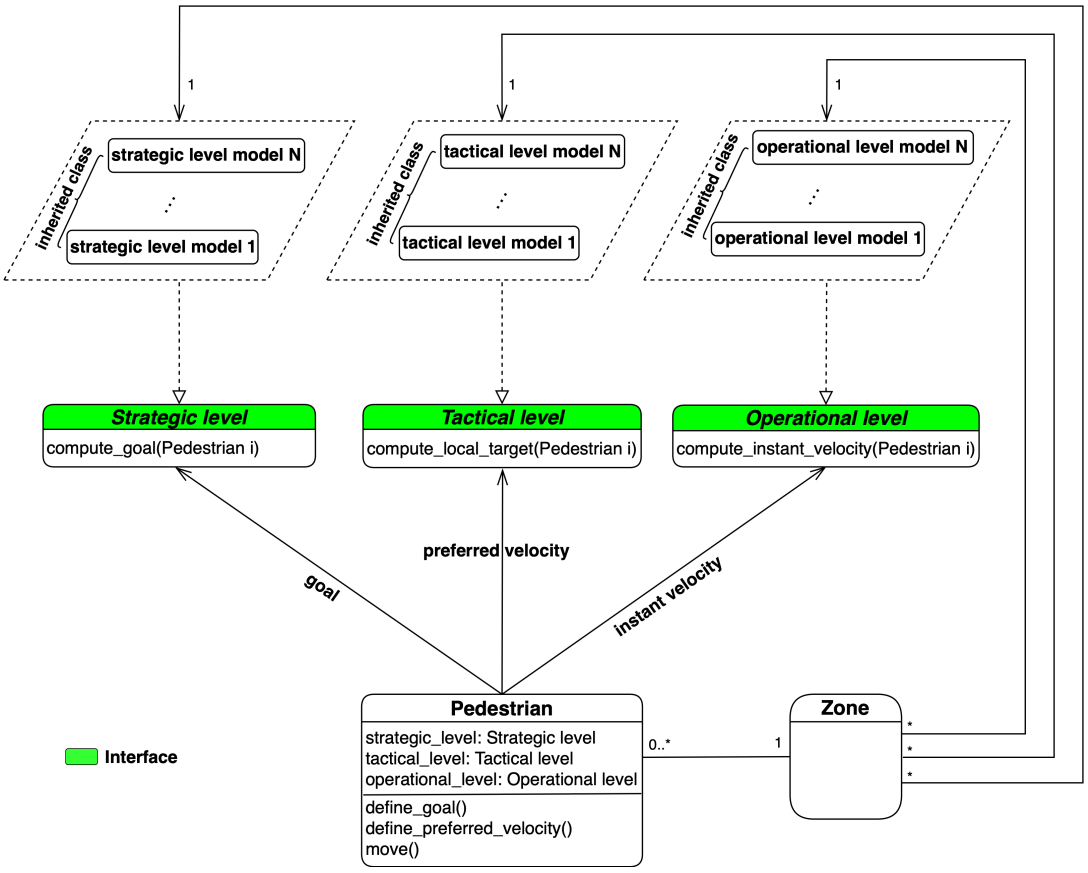


Figure 4.2: Architecture of the HyPedSim framework.

In addition, ***Zone*** class represents different areas in the environment. Each zone has a transition function responsible for monitoring when a pedestrian enters or exits the zone. When this occurs, it aggregates or disaggregates pedestrian data to update the zone. Furthermore, each zone contains information about models used to simulate pedestrian behaviors in that zone. Pedestrians update their attributes based on this information when entering a new zone to synchronize with the models used in the new zone.

Finally, it is important to note that while numerous instances of ***Pedestrian*** and ***Zone*** classes can be created in simulations, only one instance of the other classes is created (singleton design pattern). Therefore, this agent-based architecture design does not increase the complexity but enhances the flexibility in simulations by providing agents with a dynamic mechanism for switching their models at each behavioral level.

4.2.3 Pedestrian Activity Diagram

Figure 4.3 presents the activity diagram for pedestrian agents in each simulation step. Firstly, pedestrian agents perceive their surroundings to collect information about nearby agents. They then check whether they currently have a goal. If they do not, they perform *define goal* behavior, which invokes a strategic-level model inheriting from the ***Strategic level*** interface to establish a goal. After that, they determine whether they have a preferred velocity. If so, they perform *define preferred velocity* behavior which invokes a tactical-level model inheriting from ***Tactical level*** interface to compute a new preferred velocity.

Next, given a preferred velocity, pedestrian agents perform the *move* behavior, which invokes the corresponding model from the ***Operational level*** class to compute their next velocity. This velocity is then used to update the position of the pedestrian agents.

After updating their position, pedestrians determine whether they have reached their intended goal. If they have, their movement concludes. Otherwise, they check if they have entered a new zone. If so, they must update their models to align with those used in the new zone. The process involves sending a request to the new zone entity to obtain information about strategic-level, tactical-level, and operational-level models used to simulate crowd dynamics in that zone. The query returns $\mathrm{strategic\_level}_{new}$, $\mathrm{tactical\_level}_{new}$, and $\mathrm{operational\_level}_{new}$, which represent the models used in the new zone. Pedestrians then update their attributes accordingly by setting the value of the $\mathrm{strategic\_level}$ attribute to $\mathrm{strategic\_level}_{new}$, $\mathrm{tactical\_level}$ attribute to $\mathrm{tactical\_level}_{new}$, and $\mathrm{operational\_level}$ attribute to $\mathrm{operational\_level}_{new}$.

The process of perceiving the surrounding environment, defining a goal and preferred velocity, moving, and updating attributes when entering a new zone, allows pedestrian agents the ability to adapt to varying scenarios and navigate effectively in complex environments.

4.3 Application to the Festival of Lights

This section examines the HyPedSim framework's capability to model pedestrians that dynamically switch operational-level models through a case study of simulating the movement of large, dense crowds at the Festival of Lights in Lyon, France. The simulation scenario is initially described, followed by a description of how different operational-level models are combined to simulate this scenario. Next, the model combination is calibrated using a genetic algorithm that uses the real-world pedestrian outflow data collected during the festival. Finally, the results of the simulation are presented, along with a sensitivity analysis for each parameter.

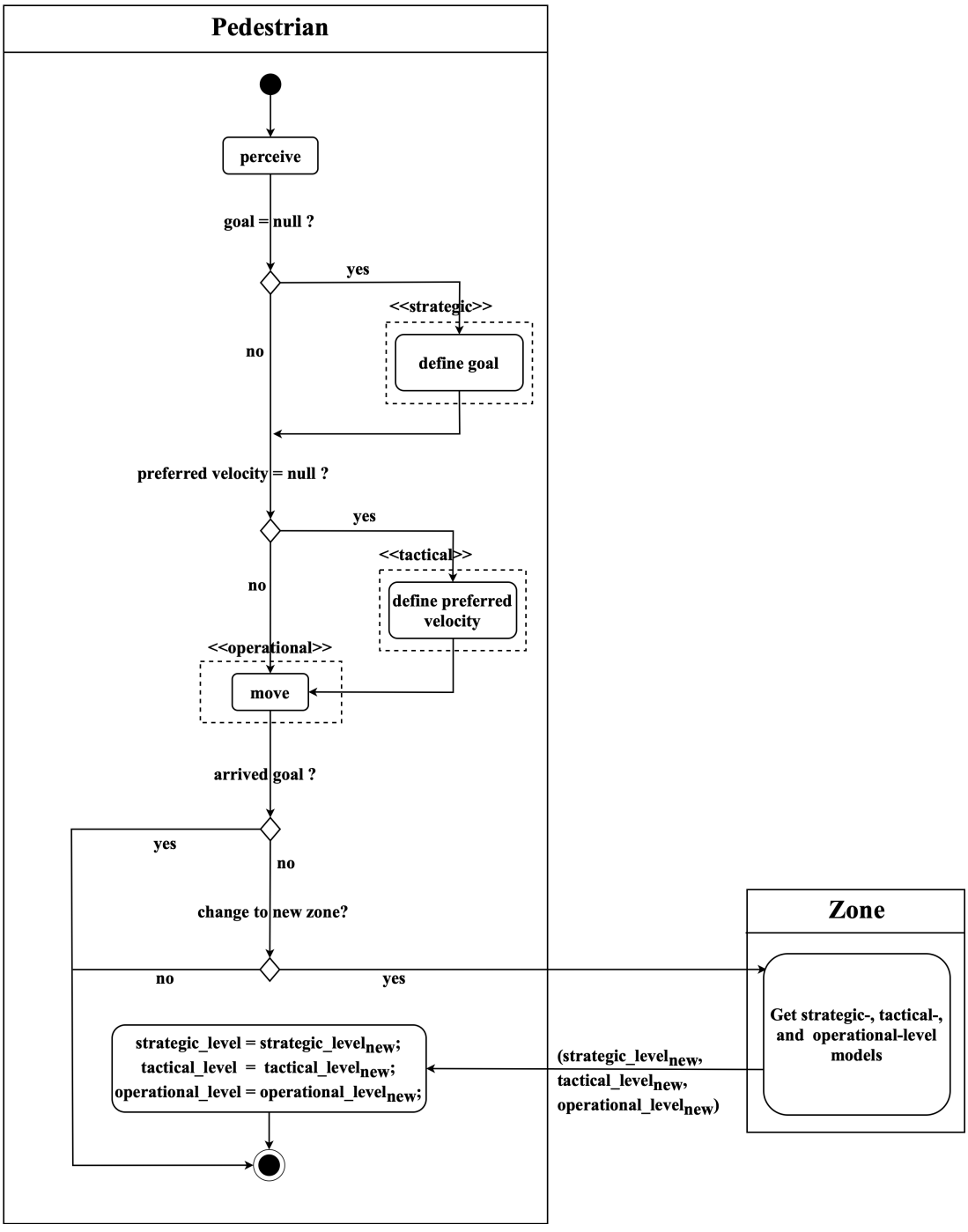


Figure 4.3: Pedestrian activity diagram at each simulation step.

4.3.1 Simulation scenario

In the previous chapter, we provided a detailed description of the Festival of Lights and of the collected empirical data during the event. The HyPedSim framework is applied to simulate large, dense crowd dynamics at Place des Terreaux, which is the main and busiest area of the festival. Figure 4.4 shows a screenshot of pedestrians gathering at Place des Terreaux to watch a projection show during the festival. Once the show ends, pedestrians move towards the two exit roads, Constantine Road and Chenavard Road, to explore other parts of the festival. We have observed from videos and on site that this creates high-density flows inside the plaza while lower-density flows form along the exit roads. We want to investigate whether the HyPedSim framework can effectively model such crowd exit scenarios where crowd density varies by location and over time. Details of the implementation are described in the next section.

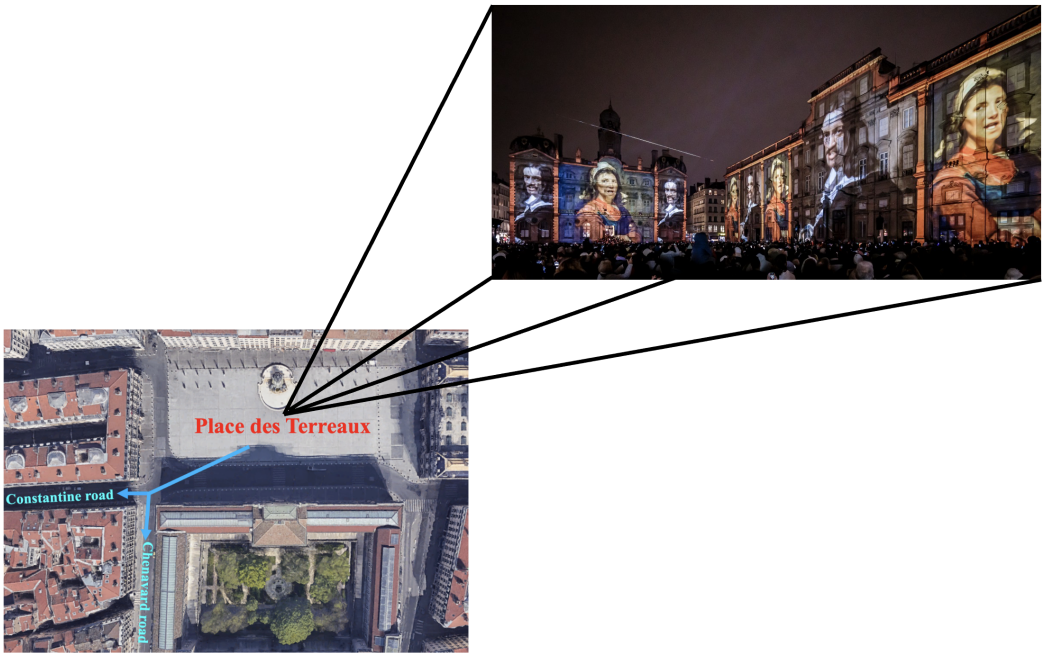


Figure 4.4: A screenshot of pedestrians watching the show at Place des Terreaux in 2022[8] and main circulation of the exiting crowd.

[8] https://www.fetedeslumieres.lyon.fr/fr/oeuvre/grand-mix-au-musee-des-beaux-arts-de-lyon (accessed on 30 January 2024)

4.3.2 Pedestrian Simulation Models

Two operational-level models are integrated into the framework: the Social Force Model [13] for simulating low-density crowds and the Continuum Crowd model [15] for simulating high-density crowds. The description of the models and their coupling to simulate a crowd exit case study at the Festival of Lights in Lyon, France, is provided in the following sections.

Social Force Model

This section describes the simplified Social Force Model (SFM) [13] integrated into the HyPedSim framework. The SFM simulates pedestrian behavior based on the assumption that pedestrian movement is influenced by social forces. These forces arise from both internal factors and external factors. The internal factors include pedestrians' intentions and motivations to reach predetermined targets, while external factors include repulsion from other pedestrians in the surrounding environment. The mathematical foundation of the SFM is based on the Newton's second law:

$$\mathbf{a}_i(t) = \frac{\mathbf{v}_i^{pref}(t) - \mathbf{v}_i(t)}{\tau} + \sum_{j \in N(i)} \mathbf{n}_{ij}(t) A \exp^{-\frac{d_{ij}(t)}{B}} + \sum_{\mathrm{w}} \mathbf{n}_{i\mathrm{w}}(t) A_{\mathrm{w}} \exp^{-\frac{d_{i\mathrm{w}}(t)}{B_{\mathrm{w}}}} \tag{4.6}$$

where $\mathbf{a}_i(t)$ and $\mathbf{v}_i(t)$ represent the acceleration vector and velocity vector at time t, respectively. The preferred velocity and preferred speed of pedestrian i are denoted as $\mathbf{v}_i^{pref}$ and $V^{pref} = \|\mathbf{v}^{pref}\|$, respectively. In this study, the same value of V^{pref} is applied to all pedestrians simulated by the SFM. In Equation 4.6, the first term describes the acceleration of pedestrian i from its current velocity to its preferred velocity within the reaction time τ. The second term represents the distance-based repulsion forces with respect to other pedestrians in the neighboring set $N(i)$, with A and B denoting the scale and interaction range of the repulsion forces, respectively, and:

$$d_{ij}(t) = \|\mathbf{x}_i(t) - \mathbf{x}_j(t)\| - (r_i + r_j) \;\text{ and }\; \mathbf{n}_{ij}(t) = \frac{\mathbf{x}_i(t) - \mathbf{x}_j(t)}{\|\mathbf{x}_i(t) - \mathbf{x}_j(t)\|} \tag{4.7}$$

where $\mathbf{x}_i(t)$ and r_i represent the position and radius of pedestrian i at time t, respectively. The separation distance between pedestrian i and j at time t is denoted as $d_{ij}(t)$, while $\mathbf{n}_{ij}(t)$ represents the unit vector directed from $\mathbf{x}_j(t)$ toward $\mathbf{x}_i(t)$. Figure 4.5 provides a geometric illustration of the repulsion force $\mathbf{f}_{ij}$ acting on pedestrian i due to the presence of pedestrian j.

Similarly, the third term in Equation 4.6 describes the repulsion forces to nearby walls, with $\mathbf{n}_{i\mathrm{w}}(t)$ representing the normalized vector perpendicular to wall W and $d_{i\mathrm{w}}(t)$ representing the

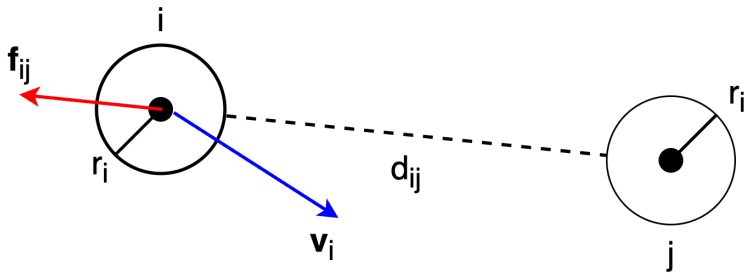


Figure 4.5: Description of repulsion force f_{ij} between pedestrian i and j.

distance to it at time t. Additionally, A_{w} and B_{w} denote the strength and interaction range of these repulsion forces.

The SFM has demonstrated effectiveness in realistically simulating well-known self-organization phenomena in pedestrian dynamics, such as lane formation in bidirectional flows and arc-shaped congestion at bottlenecks [13]. The SFM is then applied to the crowd exit case study in the Festival of Lights to simulate unidirectional pedestrian flows on long, large roads. In this scenario, the influence of the repulsion forces from walls on pedestrian movement is less significant compared to other forces acting on pedestrians, as the majority of the crowd is walking far away from walls. Furthermore, unidirectional flows create less push toward walls compared to bidirectional and multimodal flows. Hence, for simplification, the calibration mainly focuses on the first and second terms in Equation 4.6, with the parameters for repulsion forces from walls set at $A_{\mathrm{w}} = 3.0$ and $B_{\mathrm{w}} = 0.1$, based on calibrated values from previous studies [201, 202]. Based on the model description above, we aim to calibrate four parameters in the SFM, including A, B, V^{pref}, and τ.

Continuum Crowds Model

In contrast, the Continuum Crowds (CC) model, introduced by Treuille et al. [15], simulates crowd movement similarly as a continuous flow while disregarding individual-level interactions and differences. This model employs a dynamic potential field for global navigation through the eikonal equation as follows:

$$\|\nabla\phi(\mathbf{x})\| = C \tag{4.8}$$

$$\mathbf{v} = -f(\mathbf{x},\theta)\frac{\nabla\phi(\mathbf{x})}{\|\nabla\phi(\mathbf{x})\|} \tag{4.9}$$

where ϕ is the potential function and C represents the estimated unit cost along the gradient direction $\nabla\phi$. The model calculates the cost to reach the goal by considering multiple factors including spatial distance, time, and potential discomfort. Pedestrian velocity is computed so that its direction is opposite to the gradient, and its magnitude is set equal to the speed field $f(\mathbf{x}, \theta)$ evaluated at position $\mathbf{x}$ with the moving direction θ.

The CC model is typically used for simulating high-density crowds on a large scale while maintaining computational efficiency and individual agents. In this model, the environment is divided into discrete cells. Pedestrians within the same cell, who share a common target, move at the same velocity. The velocity magnitude in each cell, assuming a flat environment without slope influence, is determined based on local density according to the following equation:

$$f(\mathbf{x}, \theta) = f_{\max} - \frac{\rho(\mathbf{x} + r\mathbf{n}_\theta) - \rho_{\min}}{\rho_{\max} - \rho_{\min}}(f_{\max} - f_{flow}(\mathbf{x}, \theta)) \tag{4.10}$$

where $\rho_{\max}$ and $\rho_{\min}$ correspond to the maximum and minimum density thresholds influencing the speed value, respectively. Additionally, $f_{flow}(\mathbf{x}, \theta)$ denotes the average flow speed in the direction θ at the position $\mathbf{x}$. Similarly, $\rho(\mathbf{x} + r\mathbf{n}_\theta)$ signifies the local density evaluated at the location $\mathbf{x} + r\mathbf{n}_\theta$, which is at a distance r from $\mathbf{x}$ in the direction θ.

Figure 4.6 illustrates the relationship between density and pedestrian speed presented in Equation 4.10. Pedestrian speed reaches the maximum value of $f_{\max}$ when the density is below $\rho_{\min}$. Conversely, when the density is higher than $\rho_{\max}$, pedestrian speed is set to the average speed of the crowd flow. The lower bound for the average speed of the crowd flow is denoted as $f_{\min}$. Within these density thresholds, pedestrian speed decreases linearly from the maximum to the minimum value as the density increases from $\rho_{\min}$ to $\rho_{\max}$. For the CC model, the parameters that require calibration are $f_{\min}, f_{\max}, \rho_{\min}$, and $\rho_{\max}$.

Hybrid model

In general, pedestrian agents adhere to the activity diagram described in Figure 4.3, with additional details applied to simulate a specific case study such as the exit at the Festival of Lights in Lyon, France. These details include the implementation of specific operational-level models in appropriate areas in the environment and their coupling, the switch between operational-level models, and the scenario-specific behavior of pedestrians.

Usage of operational-level models: The crowd exit scenario at Place des Terreaux is characterized by varying densities in different areas, with high-density flows in the plaza and lower densities along the exit roads. To simulate this scenario, the environment is divided into

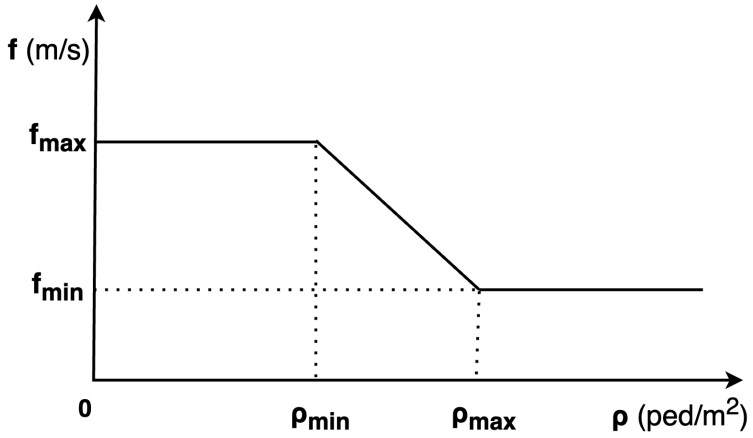


Figure 4.6: Parameters of CC model.

three separate zones as shown in Figure 4.7:

- The plaza is defined as a high-density zone and illustrated by the red rectangle.
- The two exit roads are defined as low-density zones and illustrated by the blue rectangles.

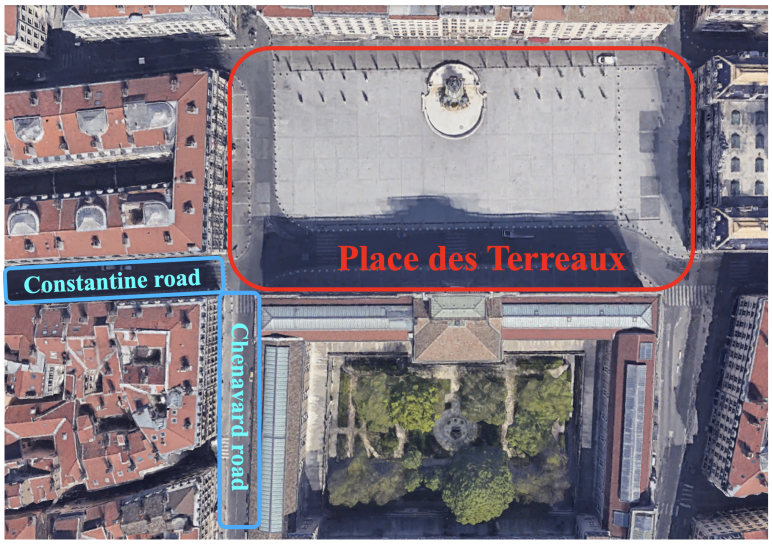


Figure 4.7: Discretization of the environment into three zones.

A specific operational-level model is executed in each zone to simulate pedestrian dynamics within that zone. The specifications of the operational-level models for these zones are as follows:

- The CC model [15] is used to simulate crowd movement in the high-density zone. In this zone, the crowd moves to a single target which is the corner leading to the two exit roads. Consequently, the CC model is applied to a single pedestrian group with a common target. The high-density zone is further divided into discrete cells, each storing information about the local average density and local average pedestrian velocity of the crowd within that cell.

- The SFM [13] is used in the low-density zones to simulate pedestrians who have left the plaza to one of the two exit roads.

Note on aggregation/disaggregation: It is crucial to recognize that combining macroscopic models and microscopic models typically requires an additional step of aggregating or disaggregating individual data for a smooth transition between these models. However, our hybrid framework integrates a mesoscopic model (CC model [15]) and a microscopic model (SFM model [13]), both of them retain individual-level representation. Therefore, our framework skips this step because individual data, including location and velocity, are conserved when an individual moves to a new zone.

Switch of operational-level models: Pedestrians switch their operational-level models when they exit the high-density zone and enter one of the low-density zones. This transition of models is achieved by updating the attribute $\mathrm{operational\_level} = SFM$. In addition, a variable t_{delay} is introduced to control the outflow of pedestrians moving from the high-density zone to the two low-density zones. This variable simulates a brief stationary behavior during the transition of operational-level models. However, a high value of t_{delay} can cause an unnaturally long pause in pedestrian movement. To prevent this, t_{delay} is capped at a maximum of 1 second. Furthermore, the t_{delay} variable must be calibrated using empirical outflow data to ensure realistic simulation results.

Exit choice behavior: After leaving the high-density zone, pedestrians probabilistically choose between the two low-density zones that correspond to the exit roads: Constantine Road and Chenavard Road. The probability of selecting an exit road is based on pedestrian proximity to each exit road:

- If their location is closer to Constantine Road, there is an α probability of selecting Constantine Road and $1 - \alpha$ probability of choosing Chenavard Road.

- On the other hand, if their location is closer to Chenavard Road, there is a β probability of selecting Chenavard Road and $1 - \beta$ probability of choosing Constantine Road.

Given the description of the hybrid model above, three parameters for hybrid coupling, including t_{delay}, α, and β, must be calibrated using empirical data collected at the Festival of Lights.

4.3.3 Simulation Details

The experimental simulation of the crowd exit case study at the Festival of Lights is configured as follows:

- The simulation runs with a time step of $\Delta t = 0.1$ s and includes a total of 3883 pedestrian agents.
- Pedestrian agents are initially distributed uniformly in the high-density zone.
- The simulation stops when all pedestrians reach the end of one of the two exit roads.
- Buildings and obstacles in the simulation environment are represented using a shapefile.
- The simulations are implemented on the GAMA platform [203] and run on a M1 MacBook Pro with 32 GB of memory.

4.3.4 Model Calibration

Calibration is an essential step in fine-tuning the parameters to have realistic simulations of crowd dynamics in real-world scenarios. This section presents the calibration process of the HyPedSim framework for simulating high-density crowd exit during the Festival of Lights using the real pedestrian outflow data collected in the festival (as described in details in Chapter 3). The calibration process involves adjusting three types of parameters with value ranges of these parameters presented in Table 4.1:

- Hybridization:
  - t_{delay}: the delay time applied to pedestrian agents transitioning from the high-density zone to the low-density zones before switching their operational-level model.
  - α: the probability that an agent, when moving out of the high-density zone and closer to Constantine, chooses Constantine road as the exit road.
  - β: the probability that an agent, when moving out of the high-density zone and closer to Chenavard road, chooses Chenavard road as the exit road.
- Parameters of SFM:

  - A: the strength of repulsion forces between pedestrians.
  - B: the interaction range of repulsion forces between pedestrians.
  - V^{pref}: the preferred speed of pedestrians.
  - τ: the reaction time of pedestrians.

- Parameters of CC model:

  - $f_{\min}$: the minimum speed of pedestrian flows.
  - $f_{\max}$: the maximum speed of pedestrian flows.
  - $\rho_{\min}$: the minimum density threshold below which pedestrian flows reach their maximum value.
  - $\rho_{\max}$: the maximum density threshold above which pedestrian flows reach their minimum value.

A Genetic Algorithm (GA) method is designed to calibrate the eleven parameters of the framework. Compared to statistical methods for parameter estimation, such as maximum likelihood estimation [204] or Bayesian inference [205], the GA method offers an intuitive understanding and an easy implementation. This method is particularly useful for a large and continuous search space [206], which is the case of this calibration task. Furthermore, the actual data used for calibration is pedestrian outflow. This macroscopic data cannot be easily and directly used with statistical methods to estimate the parameters related to individual characteristics such as preferred speed.

Each solution is evaluated based on a fitness function that measures the similarity between the simulated outflow and actual outflow. The fitness function is defined as follows:

$$fitness = \frac{1}{N}\sum_{i=1}^{N}(\frac{|f_i^{obs} - f_i^{sim}|}{f_i^{obs}})_{Constantine} + (\frac{|f_i^{obs} - f_i^{sim}|}{f_i^{obs}})_{Chenavard} \tag{4.11}$$

where f^{obs} and f^{sim} represent the observed and simulated pedestrian outflow, respectively. The fitness function quantifies the mean normalized absolute error between the observed and simulated pedestrian outflow over N time points. The lower value of the fitness function indicates a better match to the real outflow. The GA is performed via the following steps:

- **Initialization**: A population of 128 individuals (or chromosomes) is initialized in which each individual represents a string of eleven genes corresponding to the eleven parameters of the framework that need to be calibrated. The value of each gene is randomly assigned

within a predefined range of values for its corresponding parameter. These ranges are determined based on commonly used settings in the relevant literature [13, 15]. Denote ϕ_i as the i^{th} gene in a chromosome, and ϕ_i^{min} and ϕ_i^{max} are the minimum and maximum possible values for ϕ_i, respectively. Table 4.1 presents the range of values for each parameter.

Table 4.1: List of parameters for calibration and their ranges of values.

Type	Parameter ϕ_i	ϕ_i^{min}	ϕ_i^{max}
Hybrid	t_{delay}	0.0	1.0
	α	0.0	1.0
	β	0.0	1.0
SFM	A	0.5	5.0
	B	0.1	0.5
	V^{pref}	0.8	1.5
	τ	0.4	0.6
CC	$f_{\min}$	0.05	0.25
	$f_{\max}$	0.8	1.6
	$\rho_{\min}$	0.05	0.5
	$\rho_{\max}$	6.0	8.0

- **Fitness evaluation**: The fitness function calculates the mean normalized absolute error between observed and simulated pedestrian outflows, as defined in Equation 4.11. An individual's fitness value is evaluated after running a simulation using its corresponding parameter inputs. The simulation is implemented on the GAMA platform [203], with the configuration details presented in the next section. A total of $N = 20$ time points are selected systematically at a regular interval of 20 seconds throughout the crowd exit process to compare observed and simulated results.

- **Selection**: This step involves selecting superior individuals with the best fitness values from the current population to be parents for reproducing the next generation. This study applies a selection rate of 50% to choose individuals with the lowest fitness in each generation for offspring creation through genetic crossover. Furthermore, these superior individuals are preserved in the next generation to maintain high-quality solutions.

- **Crossover**: Parents chosen from the previous step are combined to generate offspring which constitute 50% of the next population. The genetic material of the parents is merged using uniform crossover, where each gene in the offspring's chromosome has a 50%

probability of coming from either parent. Figure 4.8 illustrates the uniform crossover process between two parental chromosomes, where each gene of the offspring is derived from the corresponding genes (with selected one colored in red) of either the parental chromosomes.

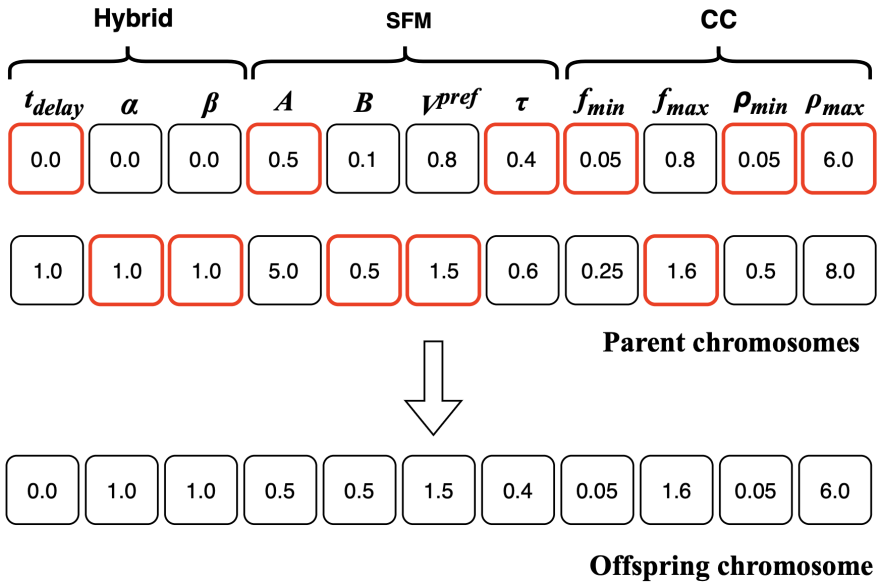


Figure 4.8: Illustration of uniform crossover process between two parental chromosomes.

- **Mutation**: Mutation is the step of making small random changes to the genes of chromosomes to maintain genetic diversity within the population. This step prevents the GA algorithm from getting stuck in a local optimum by increasing the chance of exploring new potential solutions. A mutation rate of 1% is applied to each gene in a chromosome. If a mutation occurs for a specific gene ϕ_i, the new value ϕ_i is calculated using the following formula:

$$\phi_i = \phi_i^{min} + \gamma(\phi_i^{max} - \phi_i^{min}), \quad 0 \leq \gamma \leq 1. \tag{4.12}$$

  where γ is a random value between 0 and 1.

Figure 4.9 summarizes the step-by-step calibration process presented in this study. Each iteration of the GA algorithm generates a new generation, which is then used as inputs for simulation on the GAMA platform [203]. The simulated results are compared with observed results to calculate the fitness values of each individual in this new generation. The calibration process continues until the number of generations exceeds 150.

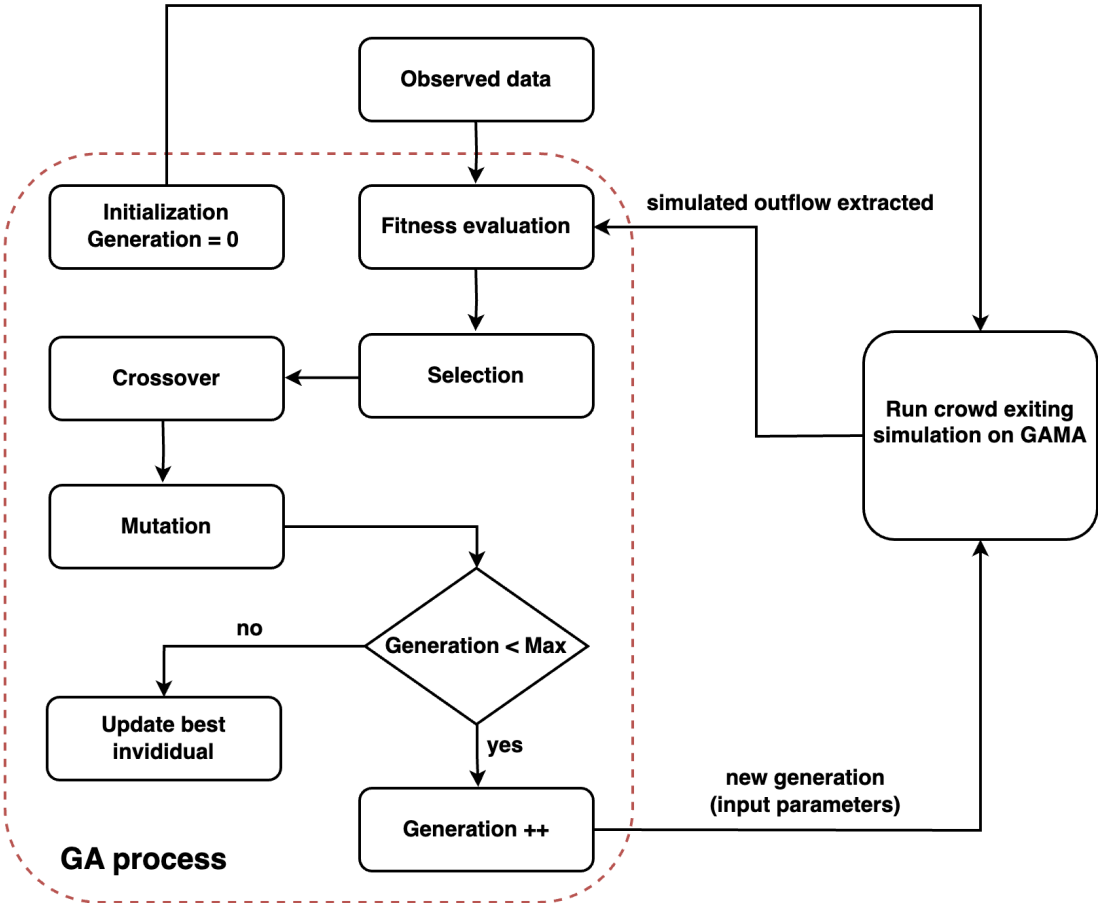


Figure 4.9: Flow of the calibration process.

4.3.5 Calibration Results

Figure 4.10 presents the average and best fitness value of the population over 150 generations. The average fitness value decreases exponentially during the first 10 generations, then more gradually over the next 25 generations, before stabilizing and remaining constant for the rest of the generations. Likewise, the best fitness value decreases rapidly in the first 30 generations and then remains stable in the remaining generations.

Table 4.2 shows the best solution obtained through calibration using the GA method over 150 generations. For hybrid parameters, the calibrated values for t_{delay}, α, and β are 0.9, 0.91, and 0.76, respectively. These values for α and β reflect the higher probability of pedestrians selecting the closest exit road. Notably, the calibrated value for α is higher than that for β, which can be attributed to the fact that more pedestrians are situated closer to Chenavard Road than

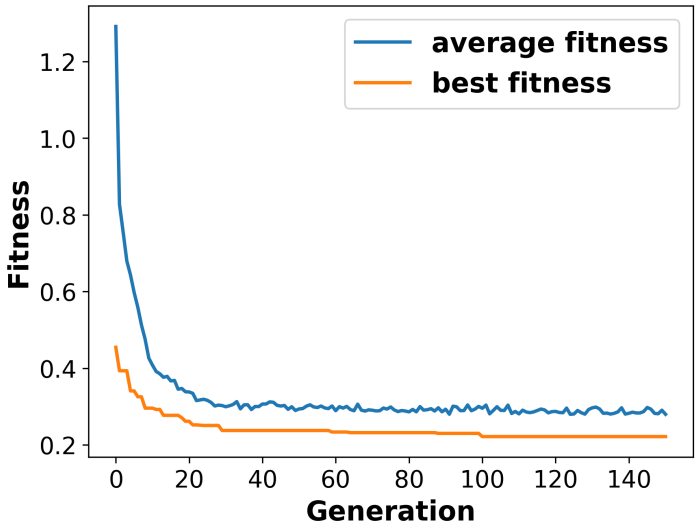


Figure 4.10: Average and best fitness values over 150 generations.

Constantine Road when they perform the exit choice behavior. Therefore, β must be lower than α to balance the outflow between the two exit roads and match the observed pedestrian outflow.

For the SFM, the parameters obtain reasonable values compared to those found in previous studies. Specifically, the calibrated values of the parameters V^{pref} and τ are $1.25\ m/s$ and $0.57s$, respectively, which align well with the mean values of the preferred speed and reaction time estimated from empirical experiments: $1.29\ m/s$ and $0.54\ s$ [202]. Similarly, the parameter A, with a value of 1.83, also falls within the range of 0.3 to 2.1 reported in [207]. However, B is 0.45, which is higher than the range of 0.18 to 0.3 found in [207].

For the CC model, the calibrates values for $f_{\min}$, $f_{\max}$, $\rho_{\min}$, and $\rho_{\max}$ are $0.15, 1.35, 0.11$ and 6.36, respectively. The relationships between the calibrated values $f_{\min}$ and $\rho_{\max}$, as well as $f_{\max}$ and $\rho_{\min}$ are qualitatively consistent with the empirical Fundamental Diagram observed in the Hajj-Tawaf mass-gathering event [208].

Table 4.2: Best solution obtained through the calibration process using the GA method.

Parameter ϕ_i	t_{delay}	α	β	A	B	V^{pref}	τ	$f_{\min}$	$f_{\max}$	$\rho_{\min}$	$\rho_{\max}$
Best value	0.9	0.91	0.76	1.83	0.45	1.25	0.57	0.15	1.35	0.11	6.36

4.3.6 Model Validation

The best-calibrated parameters obtained from calibration are then used as inputs to reproduce pedestrian outflow in the crowd exit case study at the Festival of Lights. A total of 90 simulations are run, and the average simulated pedestrian outflow is calculated with a 95% confidence interval. Figure 4.11 shows the comparison between the average simulated pedestrian outflow and observed pedestrian outflow for the two exit roads. Overall, the observed outflows exhibit sharper peaks, while the simulated outflows are smoother, yet still capture the general trends of the actual outflows. For Constantine road, the simulated outflow qualitatively captures two peaks in the actual outflow, as shown in Figure 4.11a. Conversely, for Chenevard road, the simulated outflow tends to smooth out rapid variations in the actual outflow, resulting in a more averaged representation, as displayed in Figure 4.11b.

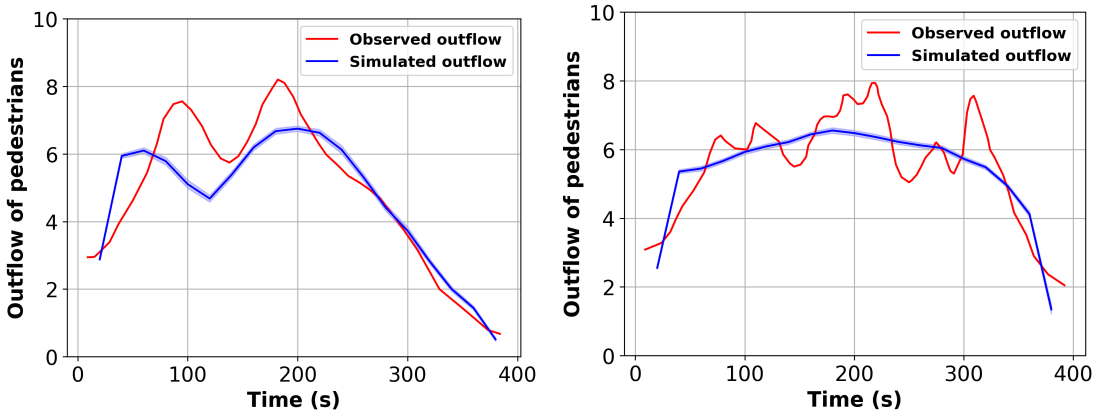


Figure 4.11: Comparison of observed and simulated outflow for the two exit roads. (**a**) Constantine road. (**b**) Chenavard road.

4.3.7 Sensitivity Analysis

While the best values for these parameters have been determined through calibration, it is essential to understand their influence on simulation outputs. Therefore, a local sensitivity analysis is performed to determine how changes in parameter values affect the simulation outputs. In this analysis, the value of each parameter varies around its calibrated values while keeping all other parameters constant at their calibrated values. The variation range is from -25% to 25% of its calibrated value, with 5% increments for each variation, as long as it remains within its corresponding minimum and maximum allowable values. This applies to almost all parameters,

except for t_{delay}, as its variation must depend on the simulation time step Δt. Specifically, t_{delay} varies from $0.5\ s$ to $1.0\ s$ with a step of $0.1\ s$. For each combination of parameter values, a total of 40 simulations are conducted and the average fitness value is calculated with a 95% confidence interval.

Figure 4.12 presents the changes in the fitness function corresponding to the variation of each parameter. For three hybridization parameters, the fitness function generally reaches the best value when the variation for each parameter stops at its calibrated value. Notably, the parameter β has the most significant impact on the simulation results, with the fitness value increasing rapidly when β deviates from 0.75, as shown in Figure 4.12c. Conversely, as seen in Figures 4.12a and 4.12b, variations in both t_{delay} and α also affect the fitness values, but their impact is not as strong as that of β.

On the other hand, the sensitivity analysis for the parameters of the operational-level models shows different trends between the SFM and CC model. For the SFM parameters, variations do not significantly affect the fitness value, as presented in Figures 4.12d to 4.12g. The reason for the insensitivity is related to the locations at which the outflow results are computed. Specifically, the outflow results are computed at the beginning part of each exit road, while SFM simulates pedestrian movement along these roads. Therefore, the initial pedestrian flow is less sensitive to these parameter changes.

For the CC parameters, only those related to pedestrian speed, including $f_{\min}$ and $f_{\max}$, show strong sensitivity to the fitness value, as seen in Figures 4.12h and 4.12i. These speed-related parameters cause a sharp increase in the fitness value when they deviate from their calibrated values. Conversely, the density-related parameters, including $\rho_{\min}$ and $\rho_{\max}$, exhibit negligible influence on the fitness value, as indicated in Figures 4.12j and 4.12k. The difference between the impacts of speed-related and density-related parameters on the fitness value is because changes in $f_{\min}$ and $f_{\max}$ cause significant changes in pedestrian speed and thus directly influence how quickly pedestrians move and reach the exit roads. In contrast, changes in $\rho_{\min}$ and $\rho_{\max}$ only cause minor adjustments in pedestrian speed, leading to a negligible impact on the fitness value.

In summary, this local sensitivity analysis provides a deep understanding of the impacts of the framework's parameters on pedestrian outflow results. Key parameters with significant impacts include three hybridization parameters: t_{delay}, α, and β, as well as two parameters in the CC model: $f_{\min}$ and $f_{\max}$. Furthermore, the analysis reconfirms that the fitness function achieves the best value when the parameters are set at their calibrated values and tends to worsen if any key parameter deviates from its calibrated value. The next section presents a further demonstration of the HyPedSim framework's capabilities.

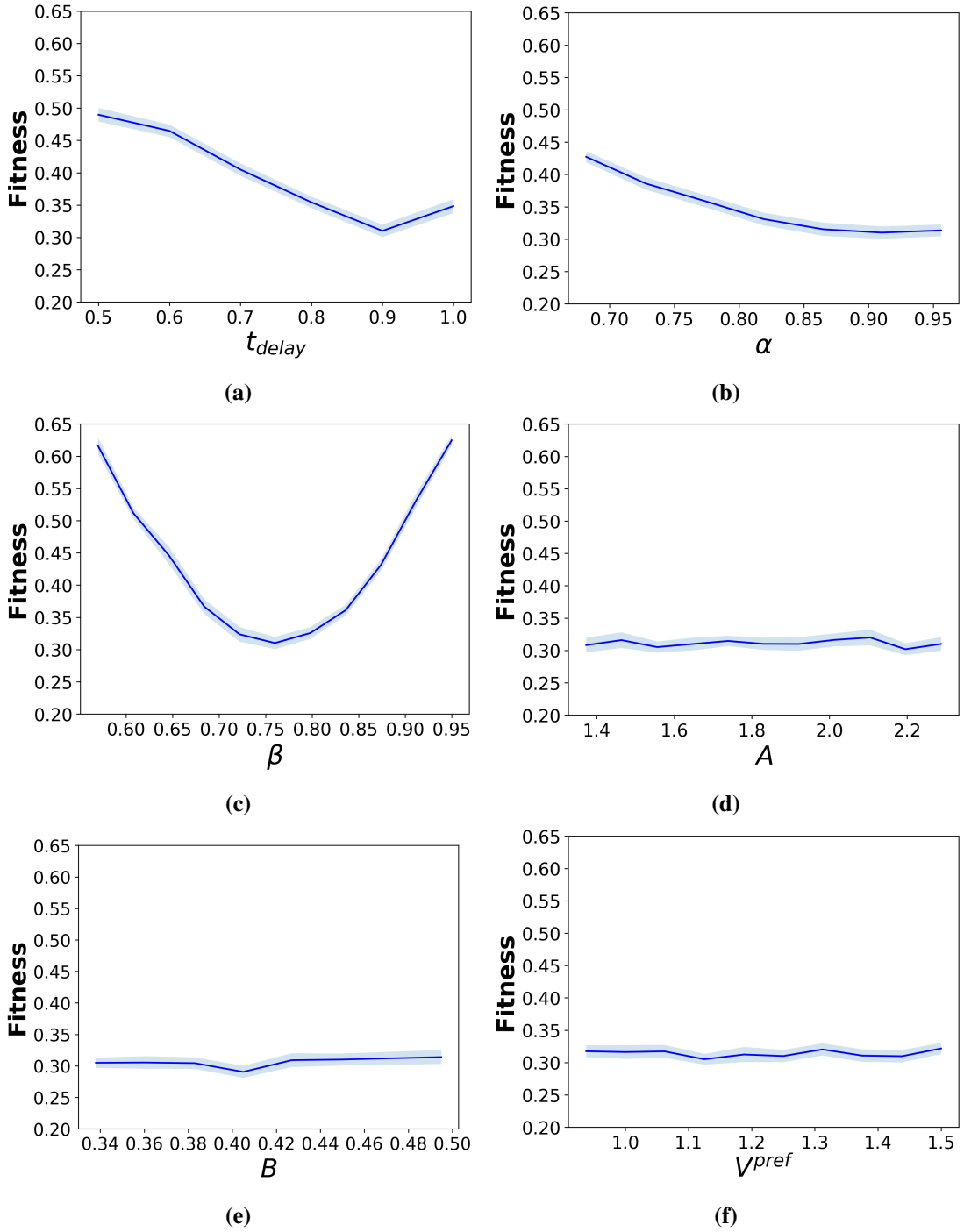


Figure 4.12: *Cont.*

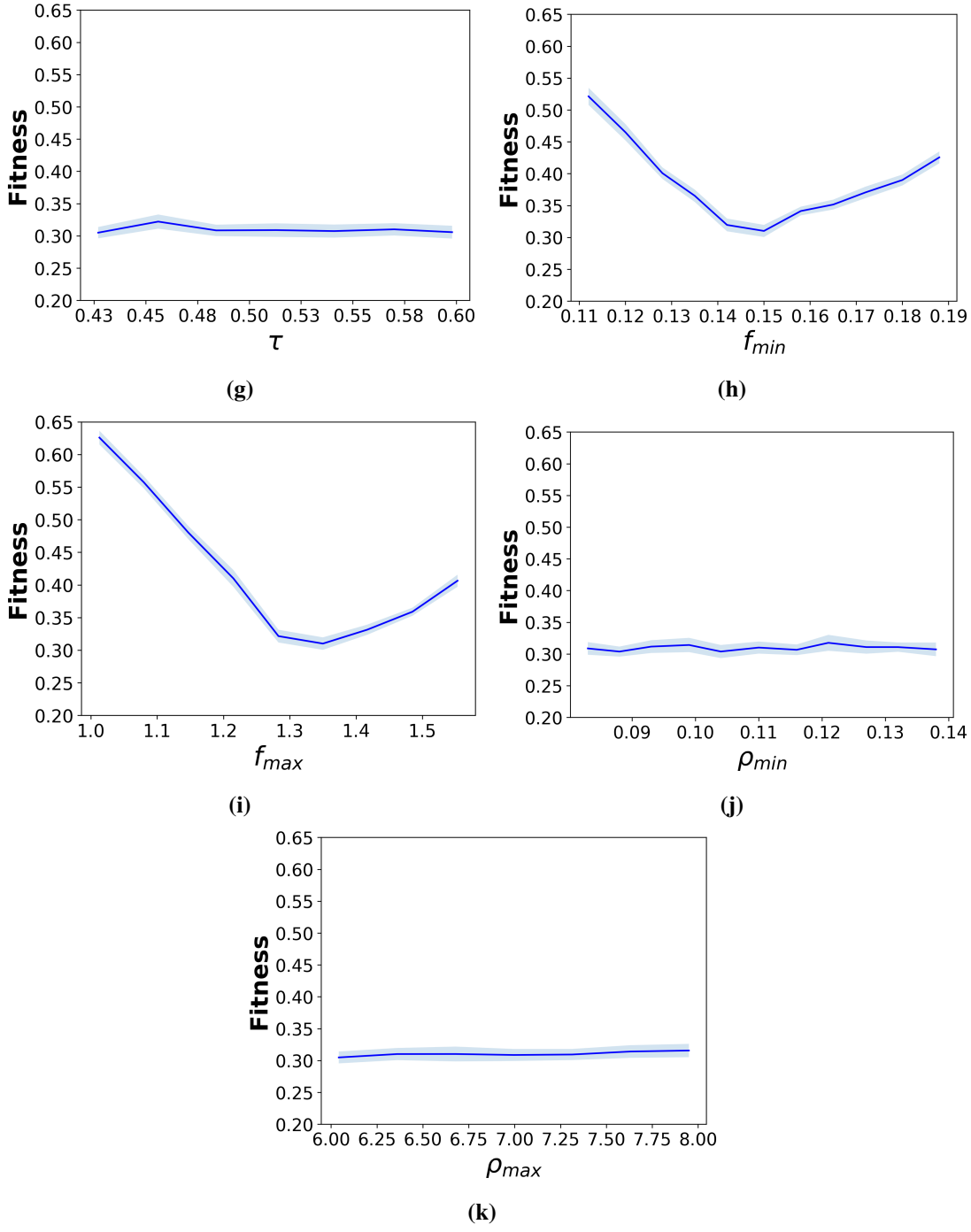


Figure 4.12: Local sensitivity analysis for different parameters. (**a**) t_{delay}, (**b**) α, (**c**) β, (**d**) A, (**e**) B, (**f**) V^{pref}, (**g**) τ, (**h**) $f_{\min}$, (**i**) $f_{\max}$, (**j**) $\rho_{\min}$, (**k**) $\rho_{\max}$.

4.4 Performance Analysis

This section demonstrates the capability of the HyPedSim framework to produce various simulations through different combinations of operational-level models. It also compares the performance of the hybrid model with these various combinations in simulating large crowds at the Festival of Lights. The flexibility of the HyPedSim framework in dynamically switching different models in each zone allows for easy implementation of these combinations. Three additional models are introduced for comparison with the proposed hybrid model:

- **SFM-only model**: This model uses only SFM to simulate pedestrian dynamics in all three zones of the environment.
- **3-CC-1 model**: This model includes three separate CC models, each simulating pedestrian dynamics in one zone. Since pedestrians in each zone move in similar directions, each CC model is designed with a single target cell. Thus, the model is named the 3-CC-1 model.
- **1-CC-2 model**: This model uses only one CC model to simulate pedestrian dynamics in the entire environment as a single zone. The CC model is designed with two different target cells to simulate two pedestrian groups heading towards the two exit roads, hence the name 1-CC-2 model. It is important to note that the complexity of the CC model increases with the number of target cells.

Each model is executed 15 times using the calibrated values for parameters to calculate the average values of the following metrics:

- Density map (measured in $\mathrm{ped}/\mathrm{m}^2$): the spatial distribution of pedestrian density over the simulated environment.
- Computation time (measured in seconds (s)): the duration required to execute one simulation step.

Figure 4.13 displays the density maps of different models simulating 6000 pedestrian agents at three specific simulation time points: t = 60 s, 150 s, and 240 s. The SFM-only model exhibits a peak density of approximately $3.5\ \mathrm{ped}/\mathrm{m}^2$ at $t = 150\ s$ observed in both the plaza and the two exit roads. Likewise, the 3-CC-1 and 1-CC-2 models show extremely high densities ranging from 6 to $8\ \mathrm{ped}/\mathrm{m}^2$, also in both the plaza and the two exit roads. In contrast, the hybrid model shows a noticeable difference in density levels between the plaza and the two exit roads. A detailed visualization of the simulation of the hybrid model is shown in Figure 4.14, where pedestrians in the plaza are simulated using the CC model and represented in red while

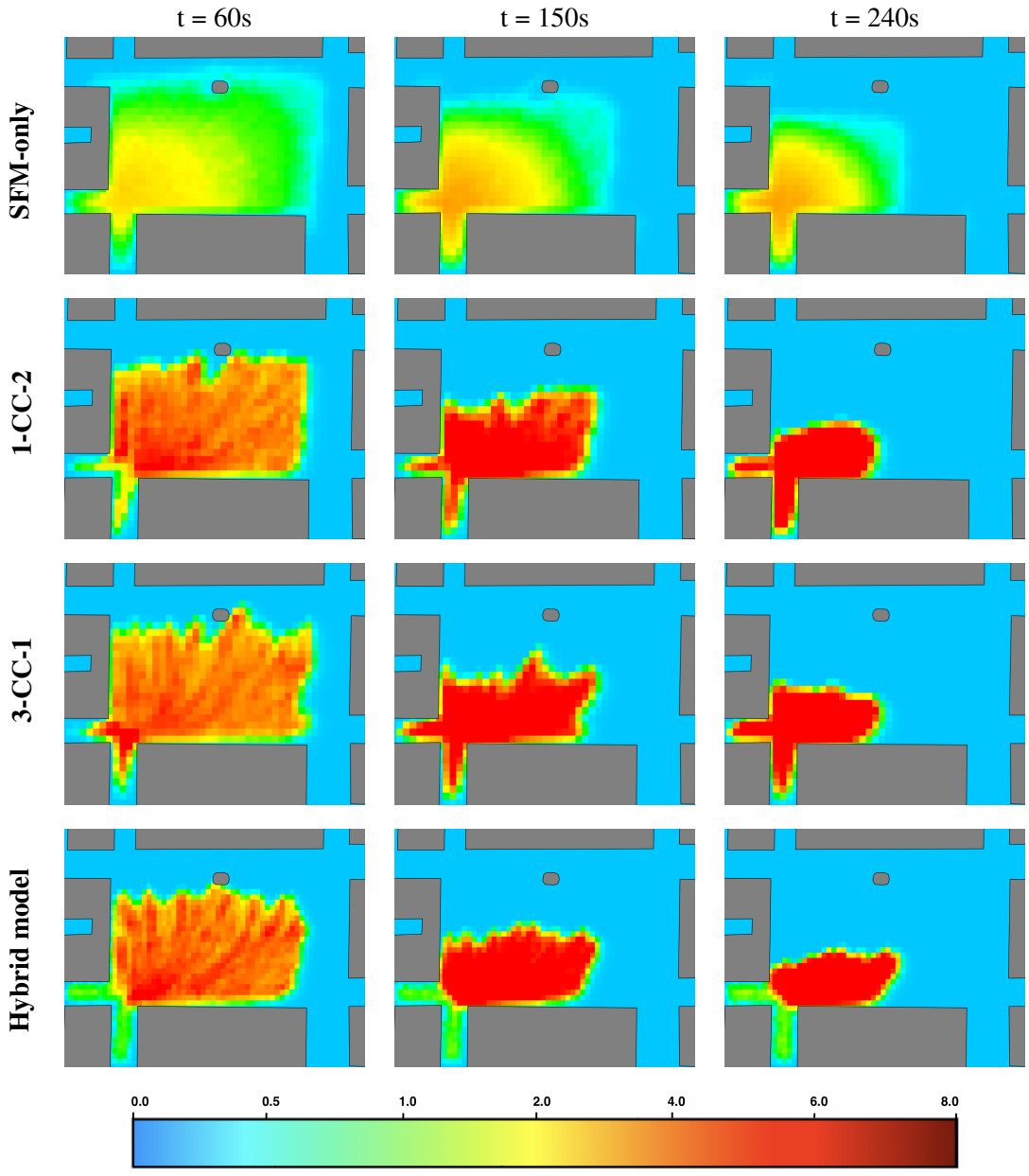


Figure 4.13: Density maps among different models with 6000 simulated agents.

pedestrians in the two exit roads are simulated using the SFM and illustrated in blue. These results suggest that the HyPedSim framework can be used to simulate various crowd scenarios by combining different zones and models. For example, the hybrid model is particularly effective in simulating pedestrian movement in environments characterized by both low and high-density areas.

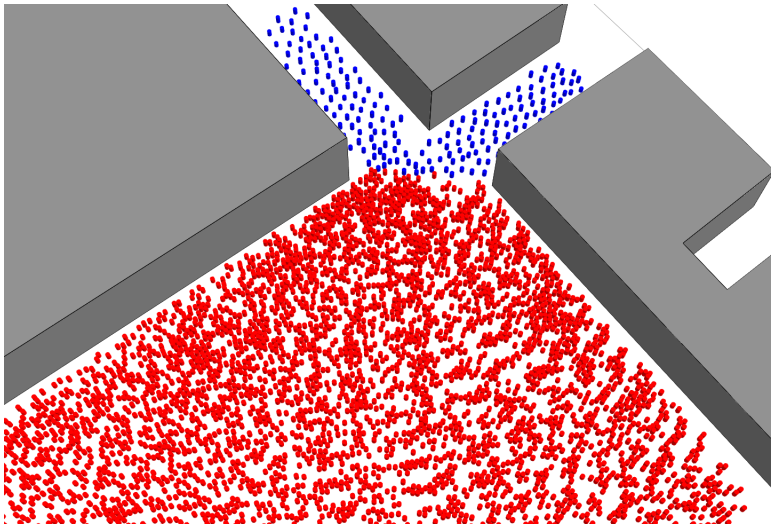


Figure 4.14: A snapshot of simulation of the hybrid model with 6000 agents.

Furthermore, each model is executed with varying numbers of agents, ranging from 3 K to 15 K, to evaluate their performance. The average duration required to perform one simulation step is computed and compared in Figure 4.15. The results indicate that the SFM-only and 3-CC-1 models take the longest time to complete each simulation step, with the duration increasing significantly as the number of agents increases. The 1-CC-2 model exhibits similar durations compared to the SFM-only and 3-CC-1 models when simulating less than 9 K agents, but it becomes faster as the number of agents exceeds 9 K. Conversely, the hybrid model demonstrates the shortest duration over different numbers of simulated agents, with a clear difference in speed compared to the other models. These results suggest that incorporating different models at each level of pedestrian behavior not only increases the ability to simulate diverse behaviors in various situations but also improves simulation performance by using appropriate models for specific areas.

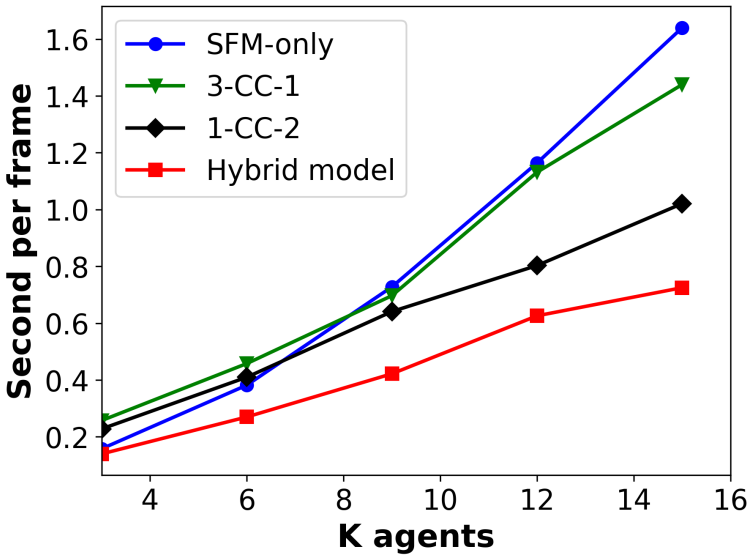


Figure 4.15: Comparison of the performance of different models.

4.5 Conclusions and Discussion

This chapter proposes an agent-based framework that allows agents to dynamically switch their operational-level models in response to changes in local density. The ability to switch between operational-level models is demonstrated through the hybrid coupling of SFM and CC model to simulate the large, dense crowd exit scenario in the Festival of Lights in Lyon, France. The simulation environment is divided into predefined zones, with the local density of each zone estimated in advance based on expert knowledge for specifying an appropriate operational-level model for each zone. The hybrid model is calibrated using a genetic algorithm that utilizes real-world pedestrian outflow data extracted from video recordings of exiting crowds at the festival. After calibration, the simulations are run with the calibrated parameters to compute the simulated outflow data. The outflow results simulated by the hybrid model can qualitatively capture the general trends of the actual outflow data. Furthermore, a local sensitivity analysis is performed to evaluate the impact of each parameter on these outflow results. The analysis shows five key parameters with significant impacts, including three hybridization parameters (t_{delay}, α, and β) and two parameters in the CC model ($f_{\min}$ and $f_{\max}$). It also reconfirms that the fitness function achieves the best value when the parameters are set at their calibrated values and tends to worsen if any of the key parameters deviates from its calibrated value.

In addition, the HyPedSim framework provides the flexibility to create various combinations of operational-level models, enabling the simulation of diverse crowd scenarios by combining different zones and operational-level models. The performance of the hybrid model is then compared to different combinations of operational-level models in terms of density map and computational time. The comparison results indicate that the hybrid model can effectively simulate pedestrians in varied density scenarios while maintaining computational efficiency compared to other model combinations. This suggests that incorporating appropriate models for specific areas not only increases the ability to simulate diverse crowd scenarios but also improves simulation performance. Moreover, the proposed agent-based architecture design is not only generic in the domain of pedestrian modeling but also extensible to other related domains, such as traffic simulation and social simulation.

However, our framework has several limitations. The first limitation is that it uses predefined, fixed zones for the simulation environment. While this approach simplifies implementation, it restricts the framework's flexibility to adapt to dynamic environments. The current approach for creating zones relies heavily on expert knowledge and specific environmental information, with no general approach applicable to any environment. Consequently, new environments may require a custom approach to zone creation, which can be time-consuming.

The second limitation is that the framework assigns models to zones based on initial density estimations, and this assignment does not change dynamically during the simulation. While it allows individual agents to dynamically switch their operational-level model when moving from one zone to another, the zones themselves cannot change their model based on real-time changes in crowd density. This means that once a zone is designated as high-density and is assigned a simulation model suitable for high-density situations, it will always use this simulation model regardless of any subsequent changes in the actual crowd density within that zone.

Future works aim to address these above limitations and enhance the adaptability of the HyPedSim framework across various scenarios. To do so, several perspectives are proposed. The first direction is to integrate a wider range of models at each behavioral level into the framework to cover a more diverse range of crowd phenomena.

In addition, dynamic zones can be used instead to address the limitations of using fixed, predetermined zones for the environment. These dynamic zones are created by clustering pedestrian coordinates using density-based clustering algorithms such as DBSCAN [209]. Each cluster then corresponds to one zone, with its boundary encompassing all pedestrians in that cluster.

Another direction is to enable each zone to change models in response to real-time changes in crowd density. To achieve this, each zone must be associated with specific triggering rules. These

rules, activated when the crowd density in the corresponding zone meets certain conditions, select an appropriate model at each behavioral level for simulating crowd dynamics within that zone. Furthermore, changes in models require an additional step for aggregating and disaggregating pedestrian data when changing from one model to another within one zone.

In the next chapter, we investigate the use of another density-related factor in hybrid approaches to improve deep learning algorithms for predicting pedestrian trajectories.

Chapter 5

TTC-SLSTM: Pedestrian Trajectory Prediction Using Time-to-collision Interaction Energy

Contents

5.1 Introduction 105
5.2 Problem formulation 106
5.3 Methodology 107
5.3.1 Time-to-collision 107
5.3.2 Interaction energy 110
5.3.3 Social-LSTM 110
5.3.4 TTC-SLSTM 112
5.4 Experiments 114
5.4.1 Datasets 115
5.4.2 Evaluation metrics 116
5.4.3 Implementation details 118
5.4.4 Results 118
5.5 Conclusion 125

In the previous chapter, we proposed HyPedSim – a hybrid framework for pedestrian simulation that allows pedestrian agents to dynamically switch their operational-level models based on density-related factors such as density-based zones. This simulation framework normally incorporates traditional operational-level models, which are typically knowledge-based and grounded in theoretical foundations. In the last decade, a new approach has emerged for predicting pedestrian movements using deep learning neural networks. Unlike traditional knowledge-based models, neural network models are data-driven, meaning they learn patterns directly from data without relying on predefined rules. This data-driven approach has gained significant attention and is showing promising results [28]. Therefore, particular interest is placed on this approach with concern about whether improvements can be made to neural network models for crowded situations using a hybrid method. This chapter investigates how other density-related factors can be integrated to improve recently emerged deep learning approaches for pedestrian trajectory prediction. This work has been presented at KSE 2023 conference and published in [7, 9, 10] where the author was the primary contributor as well as the second contributor.

5.1 Introduction

The last decade has witnessed the rapid emergence of data-driven deep learning approaches for predicting pedestrian trajectories [161, 28]. Empirical studies have shown that deep learning models, specifically long short-term memory (LSTM) and generative adversarial networks (GAN), outperform traditional physics-based models in terms of distance-based prediction accuracy [28]. Among these, LSTM networks are widely applied to pedestrian trajectory prediction due to their capability to handle time-series data. LSTM is dedicated to prediction from a time series, which corresponds to one trajectory and thus often omits the interaction with other agents. To tackle such issues, Alahi et al. have introduced Social-LSTM [11], which incorporates a pooling layer called social pooling. This layer aggregates hidden states of neighboring pedestrians to model the interactions between pedestrians. Further improvements have integrated contextual factors such as scene information [210], attention mechanisms [211], graph neural networks [212, 213], and heterogeneity among pedestrians [214]. Besides LSTM neural networks, generative adversarial networks such as Social-GAN [66] represent another type of neural network architecture. These neural networks contain a generator and a discriminator, allowing for the prediction of a distribution of potential future trajectories.

Neural network models like Social-LSTM and Social-GAN demonstrate superior performance in terms of distance-based accuracy compared to traditional physics-based models [71].

This superior performance is achieved because neural network models are trained using loss functions that primarily focus on optimizing the distance error between predicted and ground trajectories. However, recent studies indicate that predictions made by these models often result in many collisions, especially in high-density situations [10, 161].

In this chapter, we introduce an extended Social-LSTM (SLSTM) model by integrating a collision loss term into the training loss function to address the issue of unrealistic collision behavior in pedestrian trajectory prediction. The collision loss term is based on time-to-collision (TTC) interaction energy with neighboring pedestrians (therefore, we call our model TTC-SLSTM), a concept successfully applied to model interactions between pedestrians in previous research [215]. A weight factor $\lambda \geq 0$ is used to adjust the influence of the collision loss term. Our findings indicate that adding λ significantly reduces collisions and pedestrian overlaps in predicted trajectories. However, the effectiveness of λ varies across various datasets, showing different trends of evaluation metrics. Therefore, the optimal λ value must be determined for each specific dataset. In general, the best value λ is chosen to balance distance-based accuracy and collision reduction. These preliminary results suggest that the proposed TTC-SLSTM represents a promising hybrid approach for realistic pedestrian trajectory prediction, particularly in crowded scenarios.

The remainder of this chapter is organized as follows. Section 5.2 presents the mathematical formulation of the pedestrian trajectory prediction problem. Section 5.3 describes essential concepts and proposes a new methodology to improve the realism of neural network model predictions, especially in high-density scenarios. Next, Section 5.4 details the experiments and evaluations of the proposed methodology across various pedestrian datasets. Finally, Section 5.5 summarizes the results and provides a conclusion.

5.2 Problem formulation

The goal of neural network algorithms is to accurately predict the future trajectory of one or multiple target pedestrians over a specified time horizon, typically ranging from 3.6 to 4.8s [28]. This task can be characterized as a sequence-to-sequence prediction problem, wherein the input includes a sequence of observed positions of both the target pedestrian and neighboring individuals, while the output consists of a sequence of predicted positions for the target pedestrian. Mathematically, this task is formulated as follows:

- Let N represent the total number of pedestrians in a given scene.
- T_{obs} is the total number of observation time steps.

- T_{pred} is the total number of prediction time steps.
- The positions of N pedestrians in time step t are denoted as $\mathbf{p}(t) = \{\mathbf{p}_1^t, \mathbf{p}_2^t, ..., \mathbf{p}_N^t\}$, where $\mathbf{p}_i^t = (x_i^t, y_i^t)$ represents the xy-coordinate of the i^{th} pedestrian at time step t.
- Let $\hat{\mathbf{p}}_i^t = (\hat{x}_i^t, \hat{y}_i^t)$ represents the predicted position of the i^{th} pedestrian at time step t.
- **Input**: $\mathbf{p}(t)$ with $1 \leq t \leq T_{obs}$.
- **Output**: $\hat{\mathbf{p}}_i^t$ with $T_{obs} + 1 \leq t \leq T_{obs} + T_{pred}$.

The LSTM neural network [216] is a kind of recurrent neural network designed to handle the problem of long-term dependencies in sequence data. In the context of pedestrian trajectory prediction, LSTMs are used to predict the future positions of pedestrians based on their past movements. Specifically, given the past positions $\mathbf{p}_i^t$ with $1 \leq t \leq T_{obs}$ of pedestrian i as input, an LSTM outputs the predicted future positions $\hat{\mathbf{p}}_i^t$ with $T_{obs} + 1 \leq t \leq T_{obs} + T_{pred}$. The loss function in LSTM models typically minimizes the distance-based error between predicted and groundtruth trajectories. However, LSTMs do not account for interactions and dependencies among pedestrian movements in a shared space.

5.3 Methodology

This section proposes TTC-SLSTM, a neural network model that predicts pedestrian trajectories by incorporating local density factors through time-to-collision interaction energy. It starts by introducing key concepts used to develop TTC-SLSTM, including time-to-collision and interaction energy, followed by Social-LSTM neural work architecture. Finally, the section presents the TTC-SLSTM neural network, which is the extension of Social-LSTM by integrating these concepts.

5.3.1 Time-to-collision

The time-to-collision (TTC) is a crucial quantitative measure used in pedestrian dynamics that estimates the remaining time before two moving pedestrians collide if they keep moving with their current velocities [215]. Suppose that:

- Pedestrians are represented as disk-shaped agents, with r_i represents the radius of pedestrian i.
- $\mathbf{v}_i = (v_{x_i}, v_{y_i})$ represents the velocity of pedestrian i.

- The relative position and relative velocity of pedestrian i with respect to pedestrian j are denoted as $\mathbf{p}_{ij} = (x_i - x_j, y_i - y_j)$ and $\mathbf{v}_{ij} = (v_{x_i} - v_{x_j}, v_{y_i} - v_{y_j})$, respectively.

- τ_{ij} represents the estimated time to a potential future collision between pedestrians i and j.

If a collision between pedestrian i and j occurs at some time in the future, their estimated TTC is then computed as follows:

$$\tau_{ij} = \frac{-\mathbf{p}_{ij} \cdot \mathbf{v}_{ij} - \sqrt{(\mathbf{p}_{ij} \cdot \mathbf{v}_{ij})^2 - \|\mathbf{v}_{ij}\|^2(\|\mathbf{p}_{ij}\|^2 - (r_i + r_j)^2)}}{\|\mathbf{v}_{ij}\|^2} \tag{5.1}$$

where $\|.\|$ denotes Euclidean norm.

Proof. A collision between pedestrian i and pedestrian j occurs if a ray, starting from the point $\mathbf{p}_i$ and extending in the direction of the relative velocity $\mathbf{v}_{ij}$, intersects the circle $\mathcal{C}(\mathbf{p}_j, r_i + r_j)$ centered at the point $\mathbf{p}_j$ with a radius of $r_i + r_j$ at some time τ_{ij} in the future (see Figure 5.1 for a geometric illustration):

$$\begin{gathered}
\exists t > 0 \mid \mathbf{p}_i + t.\mathbf{v}_{ij} \in \mathcal{C}(\mathbf{p}_j, r_i + r_j) \\
\Leftrightarrow \exists t > 0 \mid \mathbf{p}_{ij} + t.\mathbf{v}_{ij} \in \mathcal{C}(0, r_i + r_j) \\
\Leftrightarrow \exists t > 0 \mid \|\mathbf{p}_{ij} + t.\mathbf{v}_{ij}\|^2 \leq (r_i + r_j)^2
\end{gathered}$$

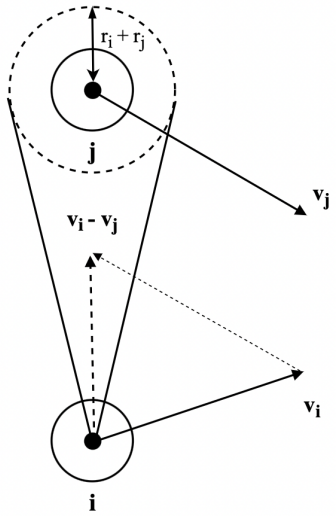


Figure 5.1: Geometric illustration of TTC calculation.

Now we need to solve this quadratic inequality by t:

$$t^2\|\mathbf{v}_{ij}\|^2 + 2t\mathbf{p}_{ij}\cdot\mathbf{v}_{ij} + \|\mathbf{p}_{ij}\|^2 - (r_i + r_j)^2 \leq 0 \qquad (\star)$$

Set $a = \|\mathbf{v}_{ij}\|^2, b = \mathbf{p}_{ij}\cdot\mathbf{v}_{ij}, c = \|\mathbf{p}_{ij}\|^2 - (r_i + r_j)^2$, and $d = b^2 - ac$.
Case 1: $d < 0$. In this case, no such t exists that satisfies the condition ($\star$). Thus, no collision will happen in the future.

Case 2: $d = 0$. Then, only $t = \dfrac{-b}{a}$ satisfies the condition ($\star$).
Case 3: $d > 0$.

- If $\dfrac{-b-\sqrt{d}}{a} \leq 0 \leq \dfrac{-b+\sqrt{d}}{a}$: pedestrians i and j are colliding.
- If $0 < \dfrac{-b-\sqrt{d}}{a} \leq \dfrac{-b+\sqrt{d}}{a}$: pedestrian i and j are going to collide in $\tau_{ij} = \dfrac{-b-\sqrt{d}}{a}$ seconds.

□

For example, consider two circular objects, A and B, each having a radius of $0.2m$, located (1, 0) and (5, 0) respectively in the 2D plane. They are moving toward each other at the velocities (1, 0) and (-1, 0), respectively, as illustrated in Figure 5.2.

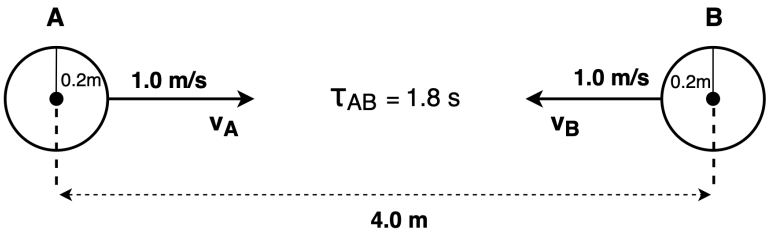


Figure 5.2: Example of computing TTC.

Then, the TTC between agents A and B is computed as follows:

$$\mathbf{p}_A = (1,0), \mathbf{p}_B = (5,0)$$
$$\mathbf{v}_A = (1,0), \mathbf{v}_B = (-1,0)$$
$$\mathbf{p}_{AB} = \mathbf{p}_A - \mathbf{p}_B = (-4,0)$$
$$\mathbf{v}_{AB} = \mathbf{v}_A - \mathbf{v}_B = (2,0)$$
$$\mathbf{p}_{AB}\cdot\mathbf{v}_{AB} = (-4,0)\cdot(2,0) = -8$$
$$\tau_{AB} = \frac{8 - \sqrt{64 - 4(16 - 0.4^2)}}{4} = 1.8(s)$$

The TTC metric is particularly useful for understanding and modeling pedestrian collision avoidance behavior. Notably, Karamouzas et al. [215] have applied this metric to investigate interactions among pedestrians. Specifically, these interactions are described by a time-to-collision interaction energy, where "interaction energy" is defined as the measure of mutual repulsion that prevents approaching pedestrians from collisions [215]. The interaction energy between pedestrians is described in the next section.

5.3.2 Interaction energy

Karamouzas et al. [215] have investigated various human motion datasets [188, 189, 217], discovering that the interaction energy of any two pedestrians is best described through a power-law function of their estimated time-to-collision, rather than just their physical distance. The interaction energy between two pedestrians, i and j, is mathematically described as:

$$E_{ij} = E(\tau_{ij}) = \frac{k}{\tau_{ij}^2} e^{-\tau_{ij}/\tau_0} \tag{5.2}$$

where k is a normalization constant, while τ_0 describes the time threshold representing that pedestrians tend to ignore potential collisions that happen in the distant future. The power-law interaction has been shown to effectively simulate various well-known crowd phenomena, such as lane formation and semi-circle shape of crowds at bottlenecks.

Figure 5.3 plots the graph of interaction energy as a function of time-to-collision, with parameter values $k = 1.5$ and $\tau_0 = 3.0$ estimated from various empirical datasets. The graph shows that the energy reaches extremely high values when τ is less than 0.5 seconds and decreases exponentially to near 0 as τ increases to 3 seconds. When τ exceeds 3 seconds, the energy is almost 0, indicating the negligible effect of potential collisions occurring in more than 3 seconds (i.e., τ_0).

5.3.3 Social-LSTM

LSTM neural networks have proven to be highly effective in learning sequence-to-sequence tasks; however, applying these neural networks for predicting pedestrian trajectories becomes more complex due to the dynamic influence of neighboring pedestrians. Unlike traditional sequence prediction tasks, pedestrian movements are significantly affected by the presence and behavior of nearby pedestrians, with the number and positions of neighboring pedestrians changing rapidly. To address this challenge, Alahi et al. [11] proposed the Social-LSTM model which incorporates a "social" pooling layer to capture the interactions between pedestrians.

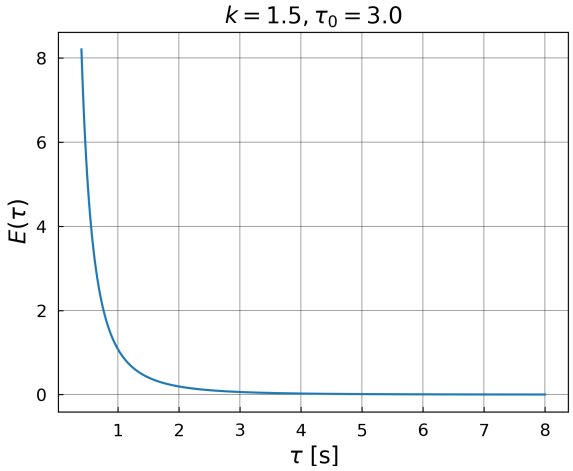


Figure 5.3: Interaction energy with $k = 1.5$ and $\tau_0 = 3.0$.

Each pedestrian trajectory in the scene is represented by an LSTM that generates a hidden state at each time step. Denote h_t^i as the hidden state of the LSTM corresponding to the ith pedestrian, which captures their latent representation at time t. The social hidden-state tensor H_i^t is created with size $N_0 \times N_0 \times D$ through grid-based pooling for the trajectory of the ith pedestrian:

$$H_i^t(m, n, :) = \sum_{j \in \mathcal{N}_i} 1_{mn}[x_j^t - x_i^t, y_j^t - y_i^t] h_j^{t-1} \tag{5.3}$$

where D and N_0 represent the hidden-state dimension and neighborhood size, respectively. In Equation 5.3, the term $1_{mn}[x, y]$ serves as the binary indicator determining whether a point (x, y) is in the (m, n) cell of the grid. Additionally, $\mathcal{N}_i$ represents the set of neighboring pedestrians within a specific radius for pedestrian i. This social pooling layer captures interactions between pedestrians by aggregating the spatial and temporal information of nearby pedestrian dynamics and sharing this information between neighboring LSTMs (as depicted in Figure 5.4).

The negative log-likelihood loss (NLL) is used to train the original SLSTM model. This loss calculates the negative logarithm of the probability that the model assigns to the true trajectory. By minimizing this loss in the training process, the model learns to assign higher probabilities to the true trajectories and lower probabilities to incorrect ones, resulting in predicted trajectories with high distance accuracy.

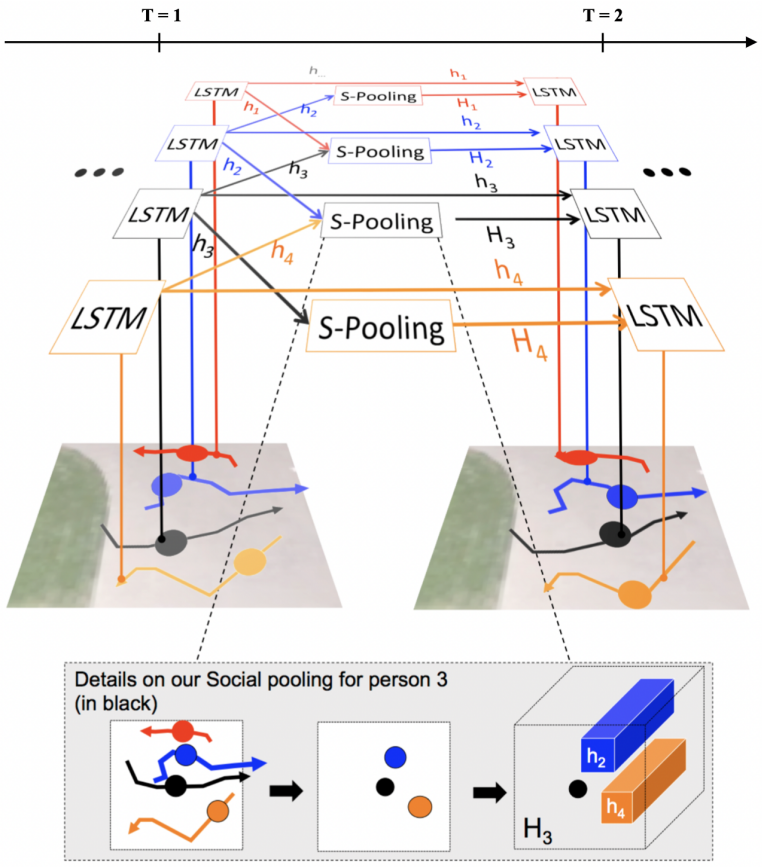
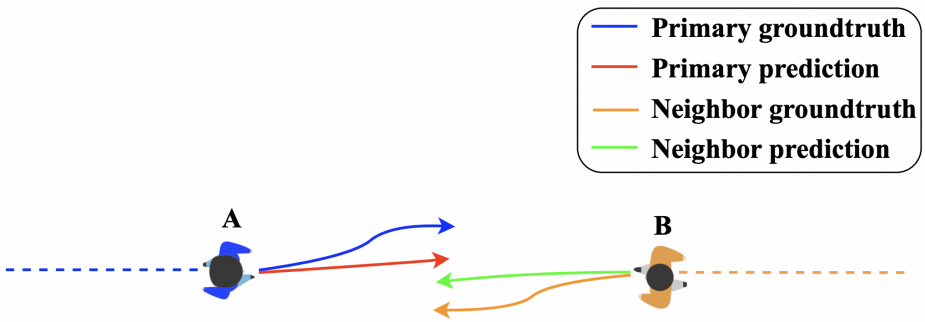


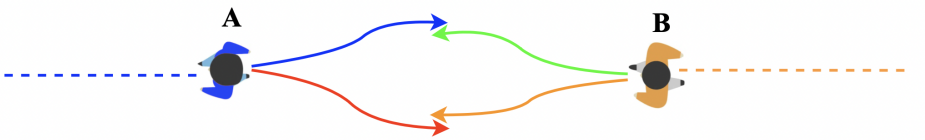
Figure 5.4: SLSTM architecture developed in [11].

5.3.4 TTC-SLSTM

Although the Social-LSTM (SLSTM) has demonstrated accurate predictions with low distance errors, the model may potentially produce trajectories that fail to reflect realistic pedestrian behaviors. More specifically, predicted trajectories can violate real-world physical constraints, resulting in physically implausible trajectories with high collision rates, especially in crowded scenarios [10]. An example illustrating this problem is presented in Figure 5.5, where two pedestrians, A and B, are moving toward each other. The groundtruth trajectories indicate that both pedestrians tend to turn to the left to avoid a collision. Figure 5.5a shows a prediction with high distance accuracy but low realism as predicted trajectories will result in a future collision. On the other hand, Figure 5.5b shows a prediction with lower distance accuracy but higher realism, where both pedestrians turn right to avoid a collision.



(a) High distance-based accuracy, low realism.



(b) Low distance-based accuracy, high realism.

Figure 5.5: Example illustrating limitation in predictions.

This issue arises because the training loss function primarily focuses on minimizing distance error without explicitly considering the physical plausibility of predicted trajectories. While the focus on distance accuracy has proven effective in low-density situations, this often results in predicted trajectories with unrealistic collisions and overlaps in crowded scenarios where there is a significant increase in interactions within the prediction time horizon [10]. To address this limitation, we propose TTC-SLSTM, a hybrid model extended from the SLSTM neural network, which incorporates physical constraints into the training loss function to improve the realism of predictions, especially in crowded situations. Specifically, the loss function is supplemented with an additional term describing a collision loss through the interaction energy with neighboring pedestrians:

$$\mathrm{L}_i = \mathrm{NLL}_i + \lambda \frac{1}{T_{pred}} \sum_{t=1}^{T_{pred}} \sum_{j \neq i} \tanh(E_{ij}) \tag{5.4}$$

where L_i and NLL_i represent the new proposed loss and the negative log-likelihood loss resulting from the i^{th} trajectory, respectively, whereas λ is a weight added to adjust the influence of the collision loss term on the total loss. If $\lambda = 0$, the model then becomes the original Social-LSTM.

The impact of different λ values on the model's prediction performance is explored in the experiments presented in the next section.

The interaction energy is chosen to represent the collision loss as it has proven to be a sufficient indicator for describing the interaction between pedestrians [215]. However, instead of directly incorporating the interaction energy function into the loss function, the $\tanh$ function, a non-linear function that has been widely used in deep learning to learn complex tasks, is applied to normalize excessively large values of interaction energy when τ is small. Figure 5.6 shows the graph of $\tanh$ function of interaction energy, with calibrated parameters $k = 1.5, \tau_0 = 3$.

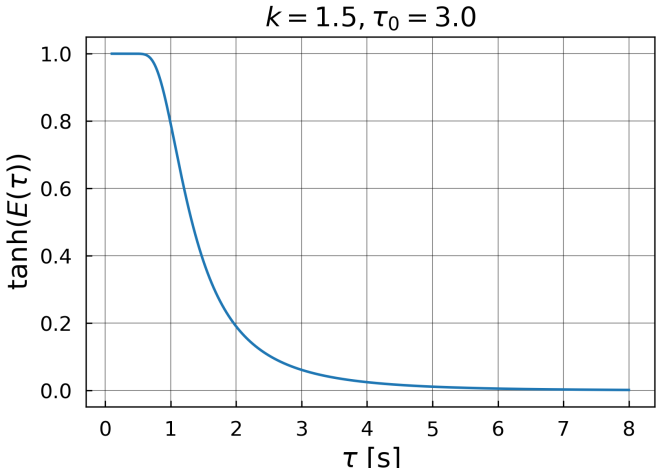


Figure 5.6: $\tanh$ function of the interaction energy $E(\tau)$ with $k = 1.5$ and $\tau_0 = 3.0$.

The second term in the new loss function (Equation 5.4) represents the total sum of the $tanh$ function applied to the energy generated when interacting with neighboring pedestrians. A penalty for collision is added to the training loss for predicted trajectories while maintaining the negative log-likelihood loss for distance error. Consequently, the neural network then learns to minimize both the negative log-likelihood loss and collision loss through the proposed training loss function.

5.4 Experiments

This presents experiments conducted to train the TTC-SLSTM on various public pedestrian datasets with different values of λ. The predictions of TTC-SLSTMs are evaluated using different metrics, including distance-based metrics (**ADE** and **FDE**) and collision metrics (**Col-I**, **Col-II**, and **AE**). Finally, the prediction results of TTC-SLSTMs are compared with those of the

SLSTM.

5.4.1 Datasets

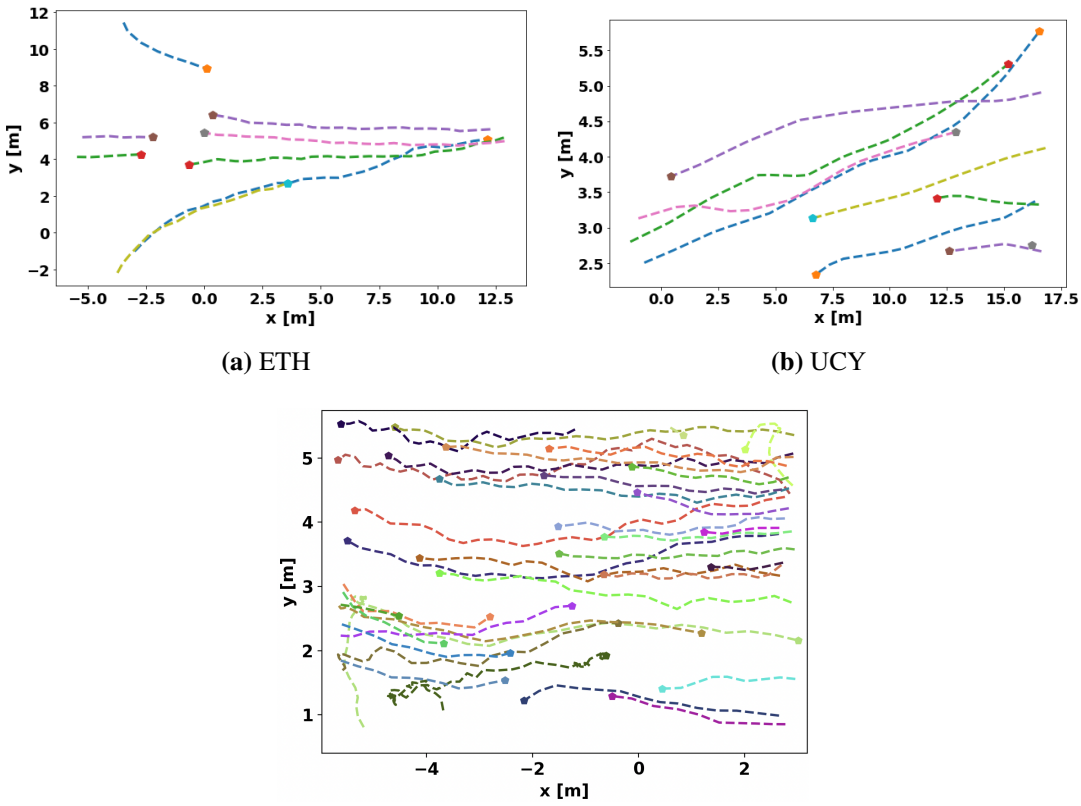
Two widely-used public pedestrian datasets, including ETH [188] and UCY [189]), together with the MADRAS dataset collected during the Festival of Lights, are selected to train the neural networks. These datasets can be categorized into two classes: low-density or high-density, based on their corresponding density level. Both low-density and high-density datasets were collected in outdoor experiments and exhibit various pedestrian traffic with unidirectional, bidirectional, and multidirectional flows. Table 5.1 summarizes the characteristics of these datasets.

Dataset	**Source**	**Data**	**Setting**	**No. Traj**	**Avg. Density** $(\mathbf{1/m^2})$	**Max. Density** $(\mathbf{1/m^2})$
Low density	ETH [188]	ETH	Outdoor	361	0.15	0.52
		HOTEL	Outdoor	389	0.13	0.32
	UCY [189]	ZARA01	Outdoor	148	0.21	0.51
		ZARA02	Outdoor	204	0.27	0.48
		UNIV	Outdoor	434	0.38	0.52
High density	MADRAS	TopView_2D	Outdoor	529	0.64	1.04
		TopView_2E	Outdoor	218	0.57	1.00
		TopView_2F	Outdoor	183	0.37	0.65

Table 5.1: Fundamental characteristics of different datasets.

For low-density datasets, the ETH dataset contains 750 trajectories and is divided into two subsets: ETH and HOTEL. Similarly, the UCY dataset includes 786 trajectories and is split into three subsets: ZARA01, ZARA02, and UCY. The average density of these low-density datasets ranges from 0.13 to 0.38 $\mathrm{ped/m^2}$, with a framerate of 2.5 fps.

For the high-density datasets, only data focusing on bidirectional flow is selected as they address more interaction between pedestrians, including TopView_2D, TopView_2E, and TopView_2F. Details about the collection and analysis of this data was presented in Chapter 3. The average density of these high-density datasets ranges from 0.37 to 0.64 $\mathrm{ped/m^2}$, with maximum density exceeding 1 $\mathrm{ped/m^2}$, with a framerate of 3.0 fps. Figure 5.7 visualizes some examples of trajectories in the UCY, ETH, and MADRAS datasets.



(a) ETH

(b) UCY

(c) MADRAS.

Figure 5.7: Trajectory examples from the datasets.

5.4.2 Evaluation metrics

Two types of metrics are used to evaluate prediction results: distance-based metrics and collision metrics. Both types are typically applied to a specific pedestrian in each prediction scene, known as the primary pedestrian.

For distance-based metrics, the **Average Displacement Error (ADE)** and **Final Displacement Error (FDE)** are used to evaluate prediction accuracy in terms of spatial similarity. ADE calculates the average Euclidean distance between the ground truth and the prediction of the primary pedestrian across all prediction time steps, providing an overall measure of how closely the predicted trajectory follows the actual trajectory throughout the entire prediction period. Conversely, FDE computes the Euclidean distance between the predicted and actual positions of the primary pedestrian at the final prediction time step. Lower values of ADE and FDE indicate higher prediction accuracy. The average ADE and FDE values for a total of M prediction scenes

are calculated as follows:

$$\begin{aligned}\textbf{ADE} &= \frac{1}{MT_{pred}} \sum_{i=1}^{M} \sum_{t=1}^{T_{pred}} \sqrt{(\hat{x}_i^t - x_i^t)^2 + (\hat{y}_i^t - y_i^t)^2} \\ \textbf{FDE} &= \frac{1}{M} \sum_{i=1}^{M} \sqrt{(\hat{x}_i^{T_{pred}} - x_i^{T_{pred}})^2 + (\hat{y}_i^{T_{pred}} - y_i^{T_{pred}})^2}\end{aligned} \tag{5.5}$$

For collision metrics, recent studies have highlighted the increasing importance of using collision metrics to evaluate the predictions of deep learning models, especially in high-density scenarios [10, 161]. Two commonly used collision metrics are **Prediction Collision (Col-I)** and **Groundtruth Collision (Col-II)** [161]. Col-I measures the percentage of prediction scenes where collisions occur between the primary pedestrian's predicted trajectories and its neighbors' predicted trajectories, indicating how often the model's predictions result in potential collisions. Similarly, Col-II calculates the proportion of prediction scenes where collisions occur between the primary pedestrian's predicted trajectories and its neighbors' actual trajectories, indicating how often the predicted trajectories collide with the actual movements of other pedestrians. These metrics provide a deeper evaluation of the realism of the predicted trajectories, particularly in high-density scenarios where there is a significant increase in collisions in predictions [9, 10]. The mathematical formulations of Col-I and Col-II are as follows:

$$\begin{aligned}\textbf{Col-I} &= \frac{1}{M} \sum_{i=1}^{M} \min(1, \sum_{j \neq i} \sum_{t=1}^{T_{pred}} [\|\hat{\mathbf{p}}_i^t, \hat{\mathbf{p}}_j^t\| < r_i + r_j]) \\ \textbf{Col-II} &= \frac{1}{M} \sum_{i=1}^{M} \min(1, \sum_{j \neq i} \sum_{t=1}^{T_{pred}} [\|\hat{\mathbf{p}}_i^t, \mathbf{p}_j^t\| < r_i + r_j])\end{aligned} \tag{5.6}$$

where $[\cdot]$ represents the Iverson bracket, with $\big[P\big]$ equaling 1 if the statement P is true and 0 otherwise. Additionally, $\|\hat{\mathbf{p}}_i^t, \hat{\mathbf{p}}_j^t\|$ denotes the Euclidean distance between the predicted positions of pedestrians i and j at time step t.

Col-I and Col-II values vary significantly with the pedestrian radius. While recent studies often use a radius of $0.1\ m$ [161], this value may not sufficiently represent the actual pedestrian body. In this study, we chose a radius of $0.2\ m$ for pedestrian representation based on the heuristic estimation of pedestrian size proposed by Moussaïd et al. [218].

Furthermore, we propose a novel collision metric for evaluating collisions in predicted pedestrian trajectories, which calculates the average energy (AE) consumed by the primary pedestrian i when interacting with other pedestrians in the corresponding prediction scene across

all prediction time steps. The average AE of M prediction scenes is computed as follows:

$$\begin{aligned} \hat{E}(\tau) &= \frac{k}{\tau^2 + \epsilon} e^{-\tau/\tau_0} \\ \textbf{AE} &= \frac{1}{MT_{pred}} \sum_{i=1}^{M} \sum_{t=1}^{T_{pred}} \sum_{j \neq i} \hat{E}^t(\tau_{ij}) \end{aligned} \tag{5.7}$$

where $\epsilon = 0.01$ is a constant added to prevent extreme values of the interaction energy $\hat{E}(\tau)$ when TTC approaches zero. While other metrics like Col-I and Col-II also measure collisions, they cannot indicate the severity of collisions in a scene, as Col-I and Col-II only provide binary values: 1 for collisions and 0 for no collisions. For example, two predicted scenes with one collision and twenty collisions would both have Col-I values of 1. Conversely, the AE metric returns a continuous value of average energy that increases with the increasing number of collisions. Thus, a predicted scene with twenty collisions would result in a higher AE value compared to a predicted scene with one collision.

5.4.3 Implementation details

The implementation follows the widely accepted settings for pedestrian trajectory prediction using neural networks, with 9 frames for observations and 12 frames for predictions [28], corresponding $T_{obs} = 9$ and $T_{pred} = 12$, respectively. For low-density datasets with fps = 2.5, it predicts a future period of $4.8\ s$ based on observations from the past $3.6\ s$. For the high-density dataset with fps = 3.0, it predicts a future period of $4.0\ s$ based on observations from the past $3.0\ s$.

The model is trained for 8 epochs with a batch size of 8, using the ADAM optimizer with a learning rate of 0.001. A hold-out validation strategy is applied, with 70% of the data used for training and 15% each for validation and testing. All calculations are performed on a M1 MacBook Pro with 32 GB of memory using PyTorch framework[9].

5.4.4 Results

The impact of the proposed collision loss term on prediction results is investigated by training the model with different values of λ ranging from 0.1 to 2.0 on the low-density and high-density datasets. This range is chosen to cover both minor and major scales of the collision loss contributing to the total loss. The training is conducted separately for each dataset type, but all

[9] `http://pytorch.org`

sub-data within each type are trained together. The prediction results of TTC-SLSTM models are analyzed and compared with those of the original SLSTM model (corresponding to $\lambda = 0.0$). For each λ, the model is trained 12 times, and the average values of various evaluation metrics, including ADE, FDE, Col-I, Col-II, and AE, are computed with a confidence interval of 95%. The results of these metrics on the low-density and high-density datasets are shown in Table 5.2 and 5.3, respectively.

ADE and FDE

For the low-density datasets, the average ADE and FDE values over the different values of λ are presented in Figures 5.8a and 5.8b, respectively. Generally, both ADE and FDE exhibit similar trends as λ increases. Slightly better ADE and FDE results compared to the baseline SLSTM model are observed when λ varies between 0.25 and 1.0. The best ADE and FDE values are achieved at $\lambda = 0.5$, with the ADE of 0.555 and the FDE of 1.135, corresponding to improvements of approximately 0.9% and 1.0%, respectively, compared to the baseline SLSTM model. However, as λ exceeds 1.0, both ADE and FDE values increase.

Likewise, the high-density datasets also exhibit continuous upward trends in average ADE and FDE results as λ increases, as displayed in Figures 5.9a and 5.9b. However, these results indicate a significantly stronger increase in ADE and FDE. The average ADE results show an increase of approximately 26.2%, from 0.416 at $\lambda = 0.0$ to 0.525 at $\lambda = 2.0$. Similarly, the average FDE results also increase around 14.2%, from 0.843 at $\lambda = 0.0$ to 0.963 at $\lambda = 2.0$.

Col-I, Col-II, and AE

For the low-density datasets, TTC-SLSTM models overperform the baseline SLSTM model for all three evaluation metrics: Col-I, Col-II, and AE, at any value of $\lambda > 0$. The average Col-I, Col-II, and AE results exhibit consistent downward trends as λ increases, with Col-I and AE results decreasing exponentially, as displayed in Figures 5.8c, 5.8d, and 5.8e, respectively. Generally, the TTC-SLSTM model with $\lambda = 2.0$ achieves the best Col-I, Col-II, and AE values across almost sub-data in the low-density datasets, as presented in Table 5.2. For Col-I, the TTC-SLSTM exhibits an exponential decline from 19.888 at $\lambda = 0.0$ to 11.689 at $\lambda = 2.0$, corresponding to an improvement of approximately 41.2%. Similarly, the average AE results show an improvement of approximately 41.7%, decreasing from 16.448 at $\lambda = 0.0$ to 9.581 at $\lambda = 2.0$. For Col-II, the average results also demonstrate a significant improvement of roughly 16.4% % from 23.619 at $\lambda = 0.0$ to 19.752 at $\lambda = 2.0$. These results suggest that a higher value of λ enhances performance for these collision metrics. However, there is always a trade-off

between the optimization of distance-error metrics (ADE, FDE) and collision metrics (Col-I, Col-II, AE). Here, the best value for λ is determined as 1.0, which results in improved ADE and Col-I metrics (see Figure 5.8f).

On the other hand, the high-density datasets only show better average Col-I, Col-II, and AE results only at small values of λ, as presented in Table 5.3. Specifically, the average Col-I and AE results decrease significantly until $\lambda = 0.5$, as displayed in Figures 5.9c and 5.9e. The average Col-I results improve by approximately 18.86% from 30.649 at $\lambda = 0.0$ to 24.868 at $\lambda = 0.5$, whereas the average AE results show an improvement of around 19.49%, from 22.686 at $\lambda = 0.0$ to 18.265 at $\lambda = 0.5$. For Col-II, the average results slightly decrease until $\lambda = 0.5$, showing an improvement of roughly 3.52%, from 45.860 at $\lambda = 0.0$ to 44.244 at $\lambda = 0.5$. However, when λ exceeds 0.5, the average values of all three collision metrics start to increase rapidly. At $\lambda = 2.0$, the average Col-I, Col-II, and AE results increase by 49.45%, 20.85%, and 30.85%, respectively, compared to their corresponding values at $\lambda = 0.0$. This may arise from the fact that high-density situations result in a significant increase in the number of pedestrians and interactions. Consequently, the total interaction energy induced by neighboring pedestrians increases substantially. Therefore, high values of λ can overly penalize the model, leading to a significantly negative impact on collision metrics. Therefore, a one-size-fits-all λ is not feasible across all datasets, and the best λ value must be defined based on each dataset's characteristics, particularly the density level. For these high-density datasets, the best λ is 0.5, as this value significantly improves collision metrics while maintaining acceptable distance-based accuracy, as seen from Figure 5.9f.

In summary, the TTC-SLSTM models show better performance than SLSTM, with a significant reduction of collisions in predicted trajectories. However, the same range of λ value affects low-density datasets and high-density datasets differently. For low-density datasets, there is a consistent decrease in collision metrics as λ increases from 0.0 to 2.0. Conversely, for high-density datasets, collision metrics decrease consistently until $\lambda = 0.5$, after which these values start to increase rapidly when λ is greater than 0.5. Therefore, the effective range for λ value must be carefully defined for each dataset.

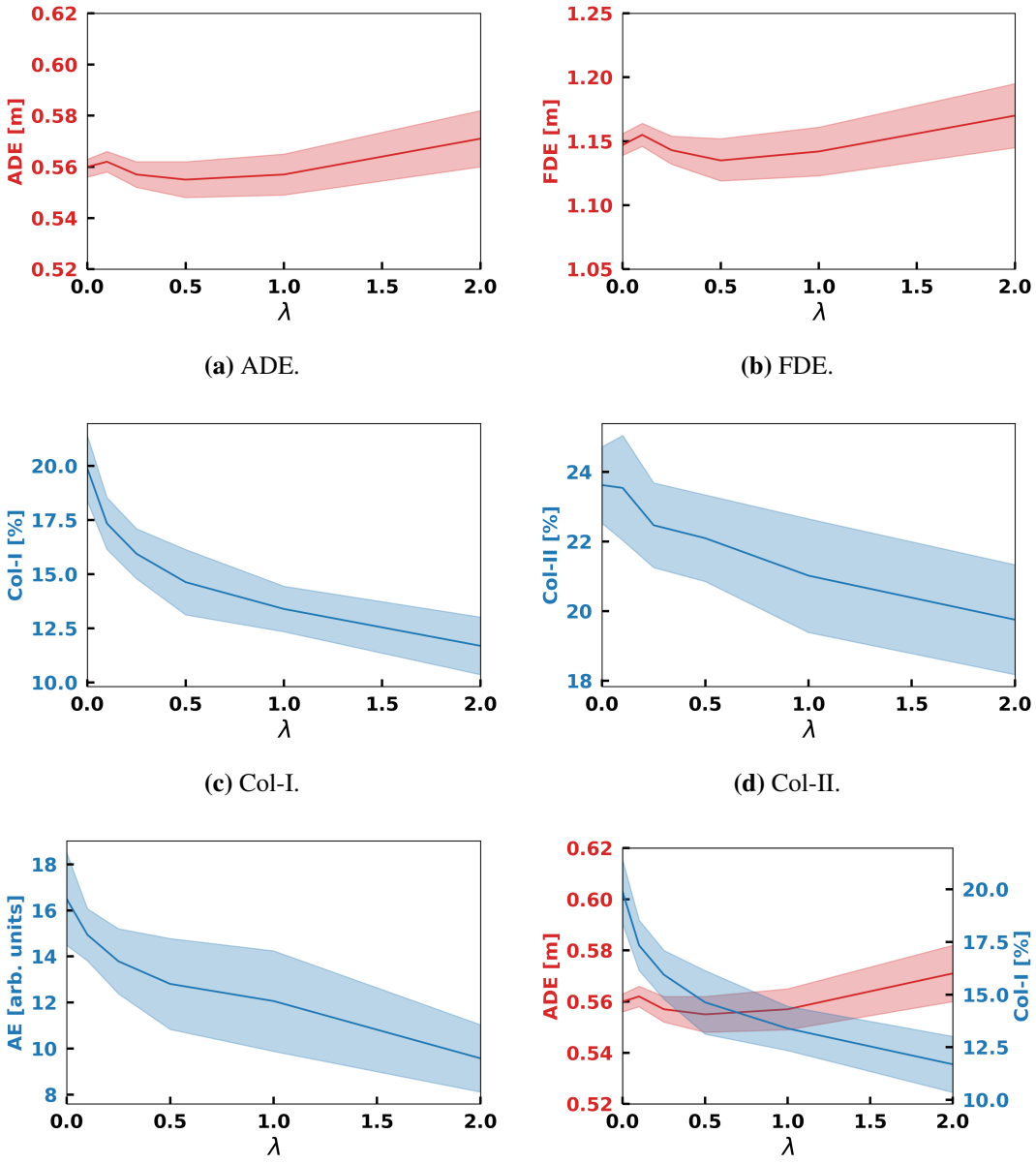
The TTC-SLSTM models can be incorporated into the HyPedSim framework at the operational level. A suitable TTC-SLSTM model can then be selected to predict pedestrian movements based on local crowd density in each zone. A TTC-SLSTM trained for low-density situations should be used in low-density zones where interactions between pedestrians are less frequent. Conversely, for high-density zones, a TTC-SLSTM model trained for high-density situations must be chosen to predict realistic pedestrian trajectories in terms of physics-based constraints.

Table 5.2: Prediction results of different models on low-density datasets.

Metric	**Data**	**SLSTM**	**TTC-SLSTM**				
			$\lambda=0.1$	$\lambda=0.25$	$\lambda=0.5$	$\lambda=1.0$	$\lambda=2.0$
ADE	ETH[188]	0.677	0.681	**0.676**	0.678	0.689	0.706
	HOTEL[188]	0.484	0.484	0.477	0.473	0.473	**0.461**
	ZARA01[189]	0.500	0.505	0.493	0.486	**0.481**	0.491
	ZARA02[189]	**0.438**	**0.438**	0.440	0.439	0.439	0.469
	UCY[189]	0.699	0.702	**0.697**	0.698	0.704	0.728
	Average	0.560	0.562	0.557	**0.555**	0.557	0.571
FDE	ETH[188]	1.302	1.309	**1.296**	1.299	1.323	1.371
	HOTEL[188]	0.882	0.890	0.878	0.865	0.875	**0.862**
	ZARA01[189]	1.049	1.065	1.042	1.017	**1.001**	1.007
	ZARA02[189]	0.949	0.952	0.956	**0.948**	**0.948**	1.004
	UCY[189]	1.555	1.558	**1.543**	1.548	1.562	1.605
	Average	1.147	1.155	1.143	**1.135**	1.142	1.170
Col-I	ETH[188]	24.568	22.828	21.661	20.218	19.855	**15.505**
	HOTEL[188]	12.145	11.430	10.130	7.379	**4.048**	4.522
	ZARA01[189]	17.381	9.285	5.972	5.476	4.998	**3.811**
	ZARA02[189]	17.978	16.822	16.752	16.126	16.281	**14.814**
	UCY[189]	27.367	26.327	25.206	23.957	21.779	**19.791**
	Average	19.888	17.338	15.944	14.631	13.392	**11.689**
Col-II	ETH[188]	28.555	29.999	29.566	28.914	28.914	**26.525**
	HOTEL[188]	17.857	16.668	14.287	13.336	**11.667**	11.668
	ZARA01[189]	30.237	29.763	28.311	25.713	21.666	**17.143**
	ZARA02[189]	19.290	**19.289**	19.612	20.060	20.678	19.752
	UCY[189]	22.158	21.970	**20.555**	22.442	22.159	23.673
	Average	23.619	23.538	22.466	22.093	21.017	**19.752**
AE	ETH[188]	20.605	17.509	17.544	13.631	13.081	**9.377**
	HOTEL[188]	11.357	11.168	9.954	9.543	6.035	**3.640**
	ZARA01[189]	10.164	6.387	3.925	**3.496**	**3.496**	6.502
	ZARA02[189]	23.186	23.456	23.036	24.387	23.943	**16.307**
	UCY[189]	17.126	16.177	14.492	12.982	13.768	**12.077**
	Average	16.488	14.939	13.790	12.808	12.065	**9.581**

Table 5.3: Prediction results of different models on high-density datasets.

Metric	Data	SLSTM	TTC-SLSTM					
			$\lambda=0.1$	$\lambda=0.25$	$\lambda=0.5$	$\lambda=1.0$	$\lambda=1.5$	$\lambda=2.0$
ADE	TopView_2D	0.440	0.442	**0.438**	0.449	0.451	0.487	0.545
	TopView_2E	**0.419**	0.434	0.426	0.440	0.455	0.481	0.549
	TopView_2F	**0.389**	0.392	0.391	0.400	0.42	0.434	0.482
	Average	**0.416**	0.423	0.418	0.430	0.442	0.467	0.525
FDE	TopView_2D	0.897	0.900	**0.892**	0.910	0.894	0.920	0.977
	TopView_2E	**0.874**	0.900	0.893	0.904	0.940	0.972	1.045
	TopView_2F	**0.758**	**0.758**	0.762	0.781	0.816	0.823	0.867
	Average	**0.843**	0.853	0.849	0.865	0.883	0.905	0.963
Col-I	TopView_2D	42.134	41.252	38.139	**35.512**	37.448	45.833	55.082
	TopView_2E	30.474	31.095	27.240	**22.512**	29.602	34.203	52.612
	TopView_2F	19.339	17.295	16.349	**16.035**	16.822	18.709	29.718
	Average	30.649	29.881	27.243	**24.686**	27.957	32.915	45.804
Col-II	TopView_2D	55.066	54.668	53.786	53.008	**52.801**	56.827	62.898
	TopView_2E	50.125	50.373	48.011	**47.017**	49.128	53.857	63.432
	TopView_2F	32.390	33.334	30.976	32.706	**30.819**	32.705	39.938
	Average	45.860	46.125	44.258	**44.244**	44.249	47.796	55.423
AE	TopView_2D	26.000	25.751	23.824	**23.136**	23.954	30.297	31.893
	TopView_2E	22.539	21.731	19.749	**18.185**	18.505	21.899	29.613
	TopView_2F	19.520	16.282	16.196	13.473	**12.404**	16.252	27.548
	Average	22.686	21.255	19.923	**18.265**	18.288	22.816	29.685



(a) ADE.

(b) FDE.

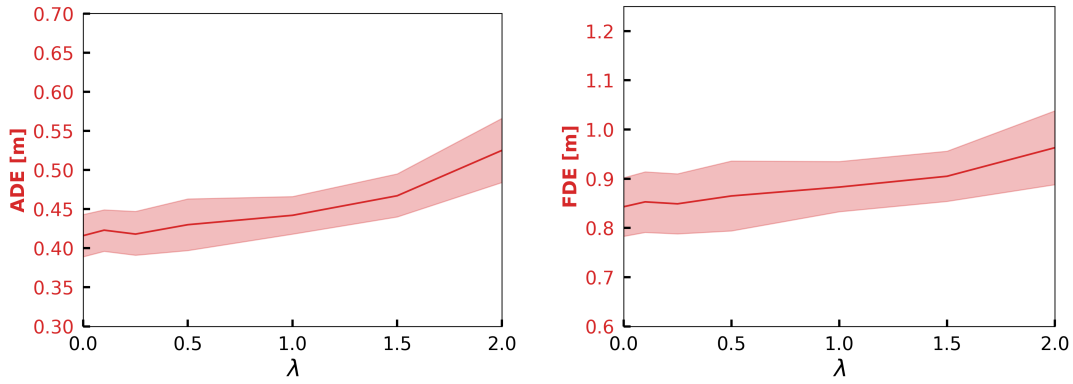
(c) Col-I.

(d) Col-II.

(e) AE.

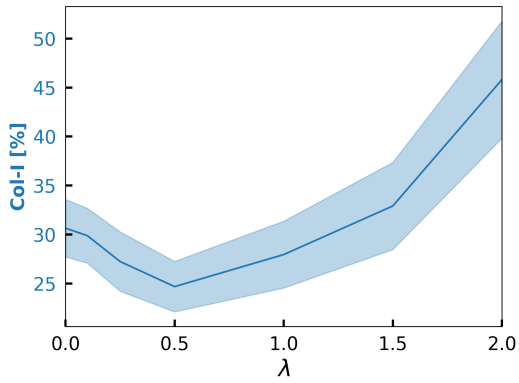
(f) ADE vs Col-I.

Figure 5.8: Average evaluation results over different values of λ on the low-density datasets.

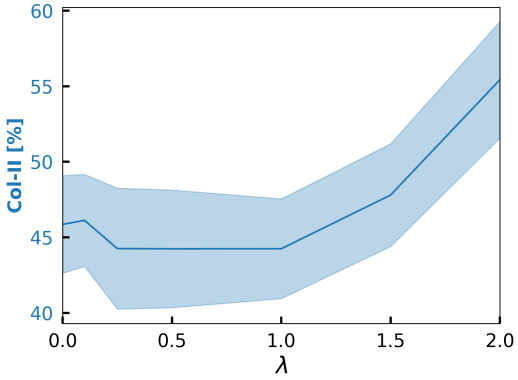


(a) ADE.

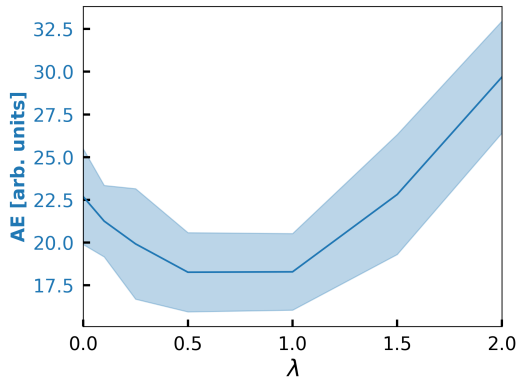
(b) FDE.



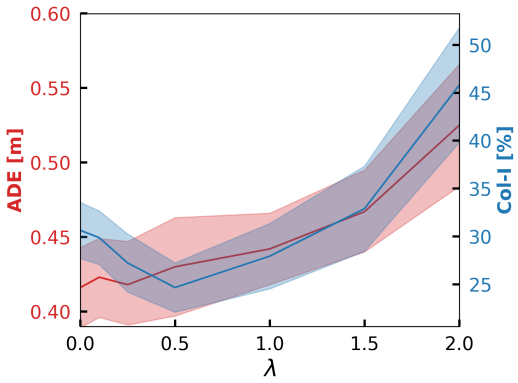
(c) Col-I.



(d) Col-II.



(e) AE.

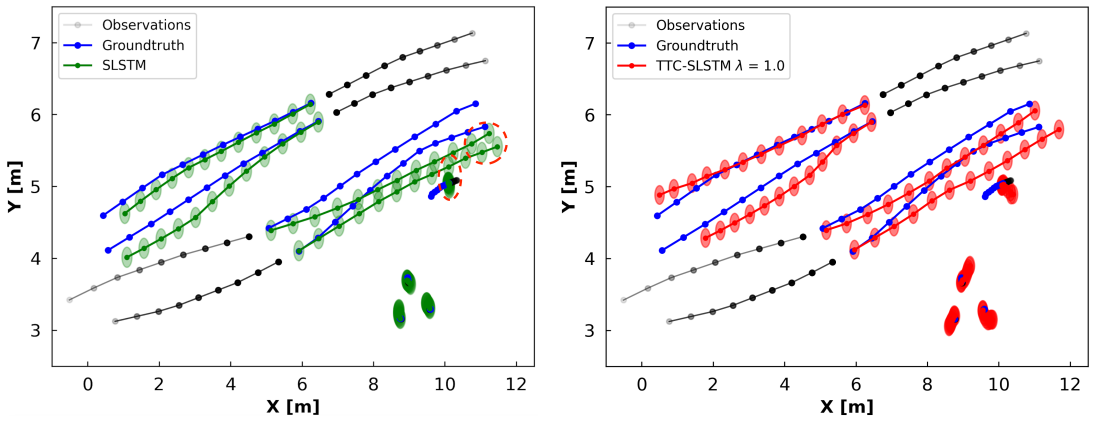


(f) ADE vs Col-I.

Figure 5.9: Average evaluation results over different values of λ on the high-density datasets.

Qualitative evaluation

Given the quantitative comparison between SLSTM and TTC-SLSTMs above, further illustrations are presented for qualitative evaluation. Figure 5.10 visually shows predictions of SLSTM and TTC-SLSTM ($\lambda = 1.0$) models for a prediction scene in the low-density datasets. There is a total of eight pedestrians in that scene, with four moving while the others are standing still. The prediction of the SLSTM model shows collisions in predicted trajectories, as highlighted by red dashed circles in Figure 5.10a. Conversely, the TTC-SLSTM model successfully handles this collision problem by generating predicted collision-free trajectories, as shown in Figure 5.10b. Furthermore, by avoiding collisions, the TTC-SLSTM model's predicted trajectories show improved distance-based accuracy compared to those generated by the SLSTM model.



(a) SLSTM. **(b)** TTC-SLSTM $\lambda = 1.0$.

Figure 5.10: Prediction examples of SLSTM and TTC-SLSTM.

5.5 Conclusion

This chapter proposes the TTC-SLSTM, an extended Social-LSTM model for pedestrian trajectory prediction that integrates a collision loss term into the training loss function to address the issue of predicted trajectories having too many collisions. The collision loss term is based on time-to-collision interaction energy with neighboring pedestrians. The impact of the collision loss term on prediction performance is investigated using various values of collision weights $\lambda \geq 0$.

The performance of the model on both low-density and high-density pedestrian datasets is evaluated using distance-based metrics (ADE and FDE) and collision metrics (Col-I, Col-II, and

AE). The prediction results indicate that adding the proposed collision loss term significantly improves the collision metrics. However, different trends in evaluation metrics are observed across these datasets. For the low-density datasets, the Col-I, Col-II, and AE results decrease exponentially while distance-based accuracy remains relatively stable. Conversely, for the high-density datasets, small λ values significantly improve the collision metrics, but high λ values overly penalize the model, leading to a negative impact with a significant increase in both distance-based and collision metrics. Therefore, the best λ value must be defined based on each dataset's characteristics, particularly the density level. These preliminary results make the proposed hybrid algorithm a promising approach for physically plausible and realistic pedestrian trajectory prediction, particularly in crowded situations.

Future work aims to achieve the following objectives for a deeper understanding of the collision loss term:

- Evaluate whether adding the collision loss term can consistently improve the realism of predictions across various neural network models beyond Social-LSTM, such as GANs and LSTM variants.
- Develop a general approach for determining the optimal λ value for each dataset based on its characteristics.
- Identify which avoidance mechanisms (e.g., steering, acceleration, and deceleration) the TTC-SLSTM learns through the use of the collision loss term.

Another promising direction to reduce collisions in predictions of neural network models is to integrate more features that aggregate contextual information into their input. Current neural network models typically use only pedestrian trajectories as inputs. These models automatically extract and learn patterns from raw trajectories provided in the input. However, several studies have shown that adding new features to the input, such as time-to-collision [5] and mean distance [219], can improve the prediction accuracy for pedestrian speed, even for simple neural networks like Multi-Layer Perceptron. To do this, the architectures of current models must be modified accordingly to accommodate additional input features.

Chapter 6

Conclusions and Discussions

Contents

6.1 Conclusions 128

6.2 Scientific contributions 130

6.3 Discussions 131

In this final chapter, we summarize the results achieved in this thesis, highlighting the scientific contributions to different research areas. We then discuss the current limitations of this study and perspectives for future work to close the thesis.

6.1 Conclusions

This thesis investigated the use of density-related factors in hybrid approaches to model and predict pedestrian movement. The investigation primarily focused on high-density situations. Based on this, four main objectives have been proposed for this study in Section 1.4, Chapter 1 summarized as follows:

- O1: review and analyze the advantages and disadvantages of modeling approaches for dense crowds. From this analysis, we identify existing gaps and propose directions to address them.
- O2: collect pedestrian data from high-density real-world experiments to calibrate and validate models for these situations.
- O3: develop an agent-based framework that combines different pedestrian simulation models to cover a broad spectrum of crowd behavior and phenomena.
- O4: improve deep learning models to predict physically realistic trajectories (i.e., no collisions) in crowded situations.

Chapter 2 addressed objective O1 by conducting a systematic review of modeling approaches for pedestrian simulation, with a particular focus on high-density situations. The review has collected a wide range of relevant articles that model pedestrian behaviors at three decisional levels: strategic, tactical, and operational. A comparative analysis was proposed to evaluate these approaches using various metrics specifically designed for high-density situations. The analysis highlighted the strengths and drawbacks of each approach and identified the following research gaps in the modeling of dense crowds:

- There is currently a lack of pedestrian datasets specifically focused on high-density scenarios in real-world settings. As a consequence, calibration and validation of models using empirical data are restricted in these situations.
- Pedestrian behaviors and crowd dynamics vary significantly with different densities. While specific models are typically used for certain scenarios, no single model can capture

the full range of pedestrian behaviors and crowd dynamics across different environments and conditions. Therefore, there is a need to integrate different models to simulate a broader spectrum of crowd phenomena.

- In recent years, deep learning neural network models have emerged and become a popular approach for predicting pedestrian trajectories. These models have shown promising results in terms of distance-based accuracy but often make unrealistic predictions with many collisions in crowded situations.

Chapter 3 presented MADRAS, a novel dataset capturing dense crowd movement. The experiment to collect this dataset was conducted in an outdoor environment during the Festival of Lights 2022 in Lyon, France, in collaboration with members of the MADRAS project. The dataset includes both macroscopic (pedestrian outflows) and microscopic (pedestrian trajectories) scales, with density up to 2.6 $\mathrm{ped}/\mathrm{m}^2$. This chapter addressed the gap in public pedestrian datasets for high-density real-world situations and thus achieved objective O2.

Chapter 4 tackled objective O3 by proposing the HyPedSim framework, a hybrid framework for pedestrian simulation. The framework allows agents to dynamically switch between different operational-level models based on predefined zones in the environment. These zones are defined according to the estimated crowd density in each area. The dynamic switching between models is demonstrated through a case study simulating dense crowds at the Festival of Lights in Lyon, France. In this case study, a hybrid coupling of the SFM and CC model is used to simulate pedestrian dynamics in different zones with varying densities. Furthermore, pedestrian outflows presented in Chapter 3 were used to calibrate the parameters of the hybrid model using a genetic algorithm. The simulated outflows of the hybrid model using calibrated parameters show that it can qualitatively capture the general trends of the actual outflow data. Additionally, a local sensitivity analysis for each parameter is conducted to provide an understanding of the impact of each parameter on the simulated outflow results.

Chapter 5 addressed objective O4 by introducing TTC-SLSTM, a novel extension of the Social-LSTM model for predicting pedestrian trajectories. This model incorporates a collision loss term in its training loss function to reduce the number of collisions in predicted trajectories. The collision loss term is derived from time-to-collision interaction energy with neighboring pedestrians. We explored the influence of the collision loss term on prediction performance using different values of collision weight $\lambda \geq 0$ across different pedestrian trajectory datasets, including both low-density (ETH [188] and UCY [189]) and high-density (MADRAS [8]). Prediction results are evaluated using distance-based metrics (ADE and FDE) and collision metrics (Col-I, Col-II, and AE). Findings reveal that the inclusion of the collision loss term

significantly reduces collisions in predicted trajectories. However, we observed different trends in evaluation metrics in these datasets. For the low-density datasets, collision metrics decrease exponentially while distance-based accuracy remains relatively stable. Conversely, for the high-density datasets, small λ values significantly improve the collision metrics, but high λ values tend to overly penalize the model, leading to a negative impact with a significant increase in both distance-based and collision metrics. These results suggest that the optimal λ should be tailored to each dataset's characteristics, particularly its density level. These preliminary results make the proposed hybrid algorithm a promising approach for physically plausible and realistic pedestrian trajectory prediction, particularly in crowded situations.

All source codes used in this study are currently being cleaned and will be shared with the public as soon as possible.

6.2 Scientific contributions

Knowledge from various research areas has been used to investigate the use of density-related factors to model and predict pedestrian movements, especially in crowded situations. In turn, the contributions of this thesis extend back to these areas, which includes three different areas: pedestrian modeling and simulation, deep learning, and agent-based modeling.

Contributions to Pedestrian modeling and simulation

This thesis makes the following three main contributions to the field of pedestrian modeling and simulation:

- A systematic review of modeling techniques for pedestrian simulation with a particular focus on high-density situations. While many reviews on pedestrian modeling and simulation are available in the literature, they specifically lack a focus on high-density situations. In comparison, our systematic review provides in-depth analyses and assessments of the strengths and drawbacks of current modeling approaches for these situations.
- A novel dataset capturing real-world high-density crowd dynamics, which is often missing in publicly available datasets. Other high-density empirical datasets are usually collected in controlled laboratory experiments where conditions are idealized, and participants know they are being observed. These factors can influence pedestrian natural behaviors, and therefore, may make the data significantly different from reality. Unlike those datasets, this new dataset was collected during the Festival of Lights in Lyon, France – a real-life mass gathering event.

- The HyPedSim framework, a hybrid framework for pedestrian simulation that allows pedestrians to dynamically switch models when moving into new zones. It offers the flexibility to create various simulations by combining different operational-level models for zones. This framework can be a valuable tool for organizers of large-scale events to test different scenarios for safety plans or evacuations.

Contributions to Deep learning

In particular, this thesis contributes TTC-SLSTM, an extended version of Social-LSTM [11], to deep learning approaches for predicting pedestrian trajectories. This contribution addresses the issue of collision-prone trajectories in neural network predictions by incorporating a collision loss term into the loss function. The prediction results of TTC-SLSTM demonstrate significant collision reduction in predicted trajectories across different pedestrian trajectory datasets. Additionally, this study proposes AE, a new collision metric that provides a continuous value to evaluate collisions, compared to discrete ones existing in the literature (i.e., Col-I and Col-II).

In general, this study contributes to deep learning by successfully demonstrating how integrating interdisciplinary knowledge improves the prediction performance of neural networks. The idea of integrating specific terms into the loss function presented in this study can be applied to other applications beyond pedestrian trajectory prediction.

Contributions to Agent-based modeling

For agent-based modeling, this thesis contributes a general agent-based architecture that allows agents to dynamically switch models based on specific conditions. In this study, agents represent pedestrians, whereas models are pedestrian simulation models. However, this architecture is not restricted to modeling pedestrian dynamics and can be extended to a wider range of applications, such as traffic simulation and emergency evacuation, to simulate environments with different kinds of dynamics.

6.3 Discussions

This study can be viewed as an initial attempt to illustrate the entire process, including reviewing modeling techniques and identifying research gaps in dense crowd simulation, collecting empirical data in high-density real-world conditions, developing models for using hybrid approaches, and validating models based on these empirical data. Although it shows significant contributions to various research domains, there exist several limitations.

For the HyPedSim framework, one of the main limitations lies in the use of predefined static zones for the environment. Although this design choice facilitates easy configuration and implementation, it restricts the ability to adapt to real-time changes in crowd density. Furthermore, this approach can create additional work when applied to a new environment, as new zones would then have to be defined. Another limitation is that once a model is assigned to a zone, it remains unchanged regardless of any variations in crowd density during the simulation. Zones cannot change their assigned operational-level models to adapt to changes in the environment. This may result in inaccurate simulations in dynamic environments. For instance, if a zone is initially expected to be high-density and is assigned a model appropriate for such situations, it will continue using this model throughout the simulation, even if the crowd in that zone disperses and the density significantly decreases.

The simulation capacity of the HyPedSim framework could be enhanced by integrating a wider range of models at each behavioral level, thereby covering a more diverse range of crowd phenomena. Furthermore, using density-based clustering algorithms like DBSCAN [209] for clustering pedestrian positions to create dynamic zones would allow for more flexible and adaptive simulations in various environments. These zones, each corresponding to one cluster, can reflect real-time changes in crowd density. Furthermore, each zone should have defined specific triggers that determine when to switch to models based on these density changes. Consequently, proper aggregation and disaggregation of pedestrian data is needed when transitioning between models to ensure the simulation accuracy.

For the TTC-SLSTM model, a deeper investigation of the impact of the loss term on prediction performance should be conducted. Firstly, the collision loss term should be tested on various neural network architectures, such as GANs or other LSTM-based models, to confirm its universal effectiveness. Next, predicted trajectories should be analyzed to understand avoidance strategies that the model learns through the collision loss term. Finally, given different trends in evaluation metrics between low-density and high-density datasets within the same range of λ, a general approach must be developed for determining the optimal value for each dataset based on its specific characteristics.

In addition, deep learning models like TTC-SLSTM can be integrated into the HyPedSim framework as operational-level models. A zone can then choose an appropriate TTC-SLSTM model based on its local density to predict pedestrian trajectories. For example, a TTC-SLSTM model trained on low-density datasets is more suitable for predicting pedestrian movement in low-density zones, where fewer interactions among pedestrians occur, and distance-based accuracy is the highest priority. In contrast, crowded situations with significant increases in pedestrian interactions require more realistic predictions that adhere to physics-based constraints.

Thus, a TTC-SLSTM model trained on high-density datasets is more appropriate for these situations.

During this thesis, I have always questioned how to combine knowledge-based simulation models and deep learning models to boost the performance of each other. Now, this has become clearer for me. Simulation models, which are built from rules and knowledge derived from domain experts, can enrich training data for deep learning models by generating synthetic data. It is particularly valuable in scenarios where real-world data is difficult or expensive to collect. Conversely, in simulations involving interactive agents, predictions of deep learning can be used to model the intelligent behaviors of these agents. Furthermore, it can be scaled up for a large number of agents through plugins allowing fast data transfer from these predictions to agents like GamPy [2]. The integration of knowledge-based and deep learning models represents a promising avenue for future research in various fields.

Glossary

- **ABM** – Agent-based Model
- **ADE** – Average Displacement Error
- **AE** – Average Energy
- **CA** – Cellular Automaton
- **CC** – Continuum Crowds
- **FD** – Fundamental Diagram
- **FDE** – Final Displacement Error
- **GA** – Genetic Algorithm
- **GAMA** – Gis & Agent-based Modelling Architecture
- **GAN** – Generative Adversarial Network
- **LSTM** – Long Short-Term Memory
- **MADRAS** – Multi-Agent modelling of dense crowd dynamics: Predict & Understand
- **NLL** – Log-likelihood Loss

- **ORCA** – Optimal Reciprocal Collision Avoidance
- **RMSE** – Root Mean Square Error
- **SFM** – Social Force Model
- **SLSTM** – Social LSTM
- **RVO** – Reciprocal Velocity Obstacle
- **TTC-SLSTM** – Time-to-Collision SLSTM
- **VO** – Velocity Obstacle

Own Bibliography

[1] Mohcine Chraibi, Jakob Cordes, Huu-Tu Dang, Oscar Dufour, Benoit Gaudou, Raphael Korbmacher, Alexandre Nicolas, David Rodney, Antoine Tordeux, and Nicolas Verstaevel. Multi-agent modelling of dense crowd dynamics (MADRAS): Application to the festival of lights (lyon). In *11th International Conference on Pedestrian and Evacuation Dynamics (PED2023)*, 2023.

[2] Huu-Tu Dang, Benoit Gaudou, and Nicolas Verstaevel. Gampy: a fast plugin for integration of python-based deep-learning models to the gama platform. In *2nd conference GAMA Days 2022*, 2022.

[3] Huu-Tu Dang, Benoit Gaudou, and Nicolas Verstaevel. A multi-level density-based crowd simulation architecture. In *International Conference on Practical Applications of Agents and Multi-Agent Systems*, pages 64–75. Springer, 2023.

[4] Huu-Tu Dang, Benoit Gaudou, and Nicolas Verstaevel. Hypedsim: A multi-level crowd-simulation framework—methodology, calibration, and validation. *Sensors*, 24(5):1639, 2024.

[5] Huu-Tu Dang, Benoit Gaudou, and Nicolas Verstaevel. Improving pedestrian dynamics predictions using neighboring factors. *Collective Dynamics*, 9:1–8, 2024.

[6] Huu-Tu Dang, Benoit Gaudou, and Nicolas Verstaevel. A literature review of dense crowd simulation. *Simulation Modelling Practice and Theory*, page 102955, 2024.

[7] Huu-Tu Dang, Raphael Korbmacher, Antoine Tordeux, Benoit Gaudou, and Nicolas Verstaevel. TTC-SLSTM: Human Trajectory Prediction Using Time-to-Collision Interaction

Energy. In *2023 15th International Conference on Knowledge and Systems Engineering (KSE)*, pages 1–6, 2023.

[8] Oscar Dufour, Huu-Tu Dang, Jakob Cordes, Raphael Korbmacher, Gaudou Benoit, Mohcine Chraibi, Alexandre Nicolas, and Antoine Tordeux. Dense crowd dynamics and pedestrian trajectories: A multiscale field study at the fête des lumières in lyon. *Scientific data*, 2014. in review.

[9] Raphael Korbmacher, Huu-Tu Dang, and Antoine Tordeux. Predicting pedestrian trajectories at different densities: A multi-criteria empirical analysis. *Physica A: Statistical Mechanics and its Applications*, 634:129440, 2024.

[10] Raphael Korbmacher, Huu-Tu Dang, Antoine Tordeux, Benoit Gaudou, and Nicolas Verstaevel. Empirical comparison of different pedestrian trajectory prediction methods at high densities. In *International Conference on Traffic and Granular Flow*, pages 231–238. Springer, 2022.

Bibliography

[11] Alexandre Alahi, Kratarth Goel, Vignesh Ramanathan, Alexandre Robicquet, Li Fei-Fei, and Silvio Savarese. Social lstm: Human trajectory prediction in crowded spaces. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 961–971, 2016.

[12] Carsten Burstedde, K Klauck, A Schadschneider, and J Zittartz. Simulation of pedestrian dynamics using a two-dimensional cellular automaton. *Physica A: Statistical Mechanics and its Applications*, 295:507–525, 06 2001.

[13] Dirk Helbing, Illés Farkas, and Tamás Vicsek. Simulating dynamic features of escape panic. *Nature*, 407:487–490, 09 2000.

[14] Roger Hughes. A continuum theory for the flow of pedestrians. *Transportation Research Part B: Methodological*, 36:507–535, 07 2002.

[15] Adrien Treuille, Seth Cooper, and Zoran Popovic. Continuum crowds. *ACM Trans. Graph.*, 25:1160–1168, 07 2006.

[16] Abhinav Golas, Rahul Narain, and Ming C Lin. Continuum modeling of crowd turbulence. *Physical review E*, 90(4):042816, 2014.

[17] Milad Haghani and Majid Sarvi. Crowd behaviour and motion: Empirical methods. *Transportation research part B: methodological*, 107:253–294, 2018.

[18] Nuria Pelechano, Jan Allbeck, and Norman Badler. Controlling individual agents in high-density crowd simulation. *Proc. Symp. Computer Animation*, 2007:99–108, 08 2007.

[19] Mingbi Zhao, Jinghui Zhong, and Wentong Cai. A role-dependent data-driven approach for high-density crowd behavior modeling. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 28(4):1–25, 2018.

[20] Dorine C Duives, Winnie Daamen, and Serge P Hoogendoorn. State-of-the-art crowd motion simulation models. *Transportation research part C: emerging technologies*, 37:193–209, 2013.

[21] Yan Feng, Dorine Duives, Winnie Daamen, and Serge Hoogendoorn. Data collection methods for studying pedestrian behaviour: A systematic review. *Building and Environment*, 187:107329, 2021.

[22] Daniel Thalmann and Soraia Raupp Musse. *Crowd simulation*. Springer Science & Business Media, 2012.

[23] Ming-Liang Xu, Hao Jiang, Xiao-Gang Jin, and Zhigang Deng. Crowd simulation and its applications: Recent advances. *Journal of Computer Science and Technology*, 29(5):799–811, 2014.

[24] Hendrik Vermuyten, Jeroen Beliën, Liesje De Boeck, Genserik Reniers, and Tony Wauters. A review of optimisation models for pedestrian evacuation and design problems. *Safety science*, 87:167–178, 2016.

[25] H Gayathri, PM Aparna, and Ashish Verma. A review of studies on understanding crowd dynamics in the context of crowd safety in mass religious gatherings. *International journal of disaster risk reduction*, 25:82–91, 2017.

[26] Dirk Helbing and Anders Johansson. *Pedestrian, Crowd and Evacuation Dynamics*, pages 6476–6495. Springer New York, New York, NY, 2009.

[27] Juliane Adrian, Nikolai Bode, Martyn Amos, Mitra Baratchi, Mira Beermann, Maik Boltes, Alessandro Corbetta, Guillaume Dezecache, John Drury, Zhijian Fu, et al. A glossary for research on human crowd dynamics. *Collective Dynamics*, 4:1–13, 2019.

[28] Raphael Korbmacher and Antoine Tordeux. Review of pedestrian trajectory prediction methods: Comparing deep learning and knowledge-based approaches. *IEEE Transactions on Intelligent Transportation Systems*, 23(12):24126–24144, 2022.

[29] Serge P Hoogendoorn and Piet HL Bovy. Pedestrian route-choice and activity scheduling theory and models. *Transportation Research Part B: Methodological*, 38(2):169–190, 2004.

[30] Francisco Martinez-Gil, Miguel Lozano, Ignacio García-Fernández, and Fernando Fernández. Modeling, evaluation, and scale on artificial pedestrians: a literature review. *ACM Computing Surveys (CSUR)*, 50(5):1–35, 2017.

[31] Dirk Helbing, Anders Johansson, and Habib Al-Abideen. Dynamics of crowd disasters: An empirical study. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 75:046109, 05 2007.

[32] Jinghong Wang, Siuming Lo, Qingsong Wang, Jinhua Sun, and Honglin Mu. Risk of large-scale evacuation based on the effectiveness of rescue strategies under different crowd densities. *Risk Analysis*, 33(8):1553–1563, 2013.

[33] Michael Wooldridge. *An introduction to multiagent systems*. John wiley & sons, 2009.

[34] Mehdi Moussaïd, Niriaska Perozo, Simon Garnier, Dirk Helbing, and Guy Theraulaz. The walking behaviour of pedestrian social groups and its impact on crowd dynamics. *PloS one*, 5(4):e10047, 2010.

[35] Lin Cheng, Ragamayi Yarlagadda, Clinton Fookes, and Prasad Yarlagadda. A review of pedestrian group dynamics and methodologies in modelling pedestrian group behaviours. *World journal of mechanical engineering*, 1(1):1–13, 2014.

[36] Serge P Hoogendoorn and Winnie Daamen. Pedestrian behavior at bottlenecks. *Transportation science*, 39(2):147–159, 2005.

[37] Wenjian Yu and Anders Johansson. Modeling crowd turbulence by many-particle simulations. *Physical review E*, 76(4):046105, 2007.

[38] Ansgar Kirchner and Andreas Schadschneider. Simulation of evacuation processes using a bionics-inspired cellular automaton model for pedestrian dynamics. *Physica A: Statistical Mechanics and its Applications*, 312:260–276, 09 2002.

[39] Suiping Zhou, Dan Chen, Wentong Cai, Linbo Luo, Malcolm Yoke Hean Low, Feng Tian, Victor Su-Han Tay, Darren Wee Sze Ong, and Benjamin D Hamilton. Crowd modeling and simulation technologies. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 20(4):1–35, 2010.

[40] Shanwen Yang, Tianrui Li, Xun Gong, Bo Peng, and Jie Hu. A review on crowd simulation and modeling. *Graphical Models*, 111:101081, 2020.

[41] Soraia Raupp Musse, Vinicius Jurinic Cassol, and Daniel Thalmann. A history of crowd simulation: the past, evolution, and new perspectives. *VISUAL COMPUTER*, 37(12, SI):3077–3092, 2021.

[42] W. van Toll and J. Pettre. Algorithms for microscopic crowd simulation: Advancements in the 2010s. *COMPUTER GRAPHICS FORUM*, 40(2):731–754, 2021.

[43] Nicola Bellomo and Christian Dogbe. On the modelling crowd dynamics from scaling to hyperbolic macroscopic models. *Mathematical Models and Methods in Applied Sciences*, 18(supp01):1317–1345, 2008.

[44] Alexandre Nicolas and Fadratul Hafinaz Hassan. Social groups in pedestrian crowds: review of their influence on the dynamics and their modelling. *Transportmetrica A: Transport Science*, 19(1):1970651, 2023.

[45] Kiran Ijaz, Shaleeza Sohail, and Sonia Hashish. A survey of latest approaches for crowd simulation and modeling using hybrid techniques. In *17th UKSIMAMSS international conference on modelling and simulation*, pages 111–116, 2015.

[46] Eleonora Papadimitriou, George Yannis, and John Golias. A critical assessment of pedestrian behaviour models. *Transportation research part F: traffic psychology and behaviour*, 12(3):242–255, 2009.

[47] Nandita Basu, Md Mazharul Haque, Mark King, Md Kamruzzaman, and Oscar Oviedo-Trespalacios. A systematic review of the factors associated with pedestrian route choice. *Transport reviews*, 42(5):672–694, 2022.

[48] Xiaomeng Shi, Zhirui Ye, Nirajan Shiwakoti, and Offer Grembek. A state-of-the-art review on empirical data collection for external governed pedestrians complex movement. *Journal of Advanced Transportation*, 2018, 2018.

[49] Almoaid Owaidah, Doina Olaru, Mohammed Bennamoun, Ferdous Sohel, and Nazim Khan. Review of modelling and simulating crowds at mass gathering events: Hajj as a case study. *Journal of Artificial Societies and Social Simulation*, 22(2), 2019.

[50] Jinghui Zhong, Dongrui Li, Zhixing Huang, Chengyu Lu, and Wentong Cai. Data-driven crowd modeling techniques: A survey. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 32(1):1–33, 2022.

[51] Fabrice Lamarche and Stéphane Donikian. Crowd of virtual humans: a new approach for real time navigation in complex and structured environments. In *Computer graphics forum*, volume 23, pages 509–518. Wiley Online Library, 2004.

[52] Sébastien Paris, Stéphane Donikian, and Nicolas Bonvalet. Environmental abstraction and path planning techniques for realistic crowd simulation. *Computer Animation and Virtual Worlds*, 17(3-4):325–335, 2006.

[53] Avneesh Sud, Russell Gayle, Erik Andersen, Stephen Guy, Ming Lin, and Dinesh Manocha. Real-time navigation of independent agents using adaptive roadmaps. In *ACM SIGGRAPH 2008 classes*, pages 1–10. ACM, 2008.

[54] Martin Stubenschrott, Christian Kogler, Thomas Matyus, and Stefan Seer. A dynamic pedestrian route choice model validated in a high density subway station. *Transportation Research Procedia*, 2:376–384, 2014.

[55] Bangquan Liu, Zhen Liu, Dechao Sun, and Chunyue Bi. An evacuation route model of crowd based on emotion and geodesic. *Mathematical Problems in Engineering*, 2018, 2018.

[56] Wouter G Van Toll, Atlas F Cook IV, and Roland Geraerts. Real-time density-based crowd simulation. *Computer Animation and Virtual Worlds*, 23(1):59–69, 2012.

[57] Xiao-Xia Jian, SC Wong, Peng Zhang, Keechoo Choi, Hong Li, and Xiaoning Zhang. Perceived cost potential field cellular automata model with an aggregated force field for pedestrian dynamics. *Transportation research part C: emerging technologies*, 42:200–210, 2014.

[58] Serge P Hoogendoorn, Femke van Wageningen-Kessels, Winnie Daamen, Dorine C Duives, and Majid Sarvi. Continuum theory for pedestrian traffic flow: Local route choice modelling and its implications. *Transportation Research Procedia*, 7:381–397, 2015.

[59] Yanqun Jiang, Bokui Chen, Xi Li, and Zhongjun Ding. Dynamic navigation field in the social force model for pedestrian evacuation. *Applied Mathematical Modelling*, 80:815–826, 2020.

[60] Yutong Sun and Hong Liu. Crowd evacuation simulation method combining the density field and social force model. *Physica A: Statistical Mechanics and its Applications*, 566:125652, 2021.

[61] Dirk Helbing and Peter Molnar. Social force model for pedestrian dynamics. *Physical Review E*, 51, 05 1995.

[62] Jur van den Berg, Ming Lin, and Dinesh Manocha. Reciprocal velocity obstacles for real-time multi-agent navigation. In *Proceedings of 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, USA, May 19-23, 2008*, pages 1928–1935, 05 2008.

[63] Jur Van Den Berg, Stephen J Guy, Ming Lin, and Dinesh Manocha. Reciprocal n-body collision avoidance. In *Robotics Research: The 14th International Symposium ISRR*, pages 3–19. Springer, 2011.

[64] Wei Shao and Demetri Terzopoulos. Autonomous pedestrians. *Graphical models*, 69(5-6):246–274, 2007.

[65] Samuel Lemercier, Asja Jelic, Richard Kulpa, Jiale Hua, Jerome Fehrenbach, Pierre Degond, Cecile Appert-Rolland, Stéphane Donikian, and Julien Pettre. Realistic following behaviors for crowd simulation. *Computer Graphics Forum*, 31:489–498, 05 2012.

[66] Agrim Gupta, Justin Johnson, Li Fei-Fei, Silvio Savarese, and Alexandre Alahi. Social gan: Socially acceptable trajectories with generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2255–2264, 2018.

[67] Rahul Narain, Abhinav Golas, Sean Curtis, and Ming C Lin. Aggregate dynamics for dense crowd simulation. In *ACM SIGGRAPH Asia 2009 papers*, pages 1–8. ACM, 2009.

[68] Muzhou Xiong, Michael Lees, Wentong Cai, Suiping Zhou, and Malcolm Yoke Hean Low. Hybrid modelling of crowd simulation. *Procedia Computer Science*, 1(1):57–65, 2010. ICCS 2010.

[69] Nguyen Thi Ngoc Anh, Zucker Jean Daniel, Nguyen Huu Du, Alexis Drogoul, and Vo Duc An. A hybrid macro-micro pedestrians evacuation model to speed up simulation in road networks. In *Advanced Agent Technology: AAMAS 2011 Workshops, AMPLE, AOSE, ARMS, DOCM 3 AS, ITMAS, Taipei, Taiwan, May 2-6, 2011. Revised Selected Papers 10*, pages 371–383. Springer, 2012.

[70] Xiaoxia Yang and Qianling Wang. Crowd hybrid model for pedestrian dynamic prediction in a corridor. *IEEE Access*, 7:95264–95273, 2019.

[71] Raphael Korbmacher and Antoine Tordeux. Review of pedestrian trajectory prediction methods: Comparing deep learning and knowledge-based approaches. *IEEE Transactions on Intelligent Transportation Systems*, 23(12):24126–24144, 2022.

[72] Sean Curtis, Andrew Best, and Dinesh Manocha. Menge: A modular framework for simulating crowd movement. *Collective Dynamics*, 1:1–40, 2016.

[73] Milad Haghani and Majid Sarvi. Human exit choice in crowded built environments: Investigating underlying behavioural differences between normal egress and emergency evacuations. *Fire Safety Journal*, 85:1–9, 2016.

[74] Hai-Jun Huang and Ren-Yong Guo. Static floor field and exit choice for pedestrian evacuation in rooms with internal obstacles and multiple exits. *Physical Review E*, 78(2):021131, 2008.

[75] Guo Ren-Yong and Huang Hai-Jun. Logit-based exit choice model of evacuation in rooms with internal obstacles and multiple exits. *Chinese physics B*, 19(3):030501, 2010.

[76] Ahmed Abdelghany, Khaled Abdelghany, Hani Mahmassani, and Wael Alhalabi. Modeling framework for optimal evacuation of large-scale crowded pedestrian facilities. *European Journal of Operational Research*, 237(3):1105–1118, 2014.

[77] Ali Edrisi, Bahman Lahoorpoor, and Ruggiero Lovreglio. Simulating metro station evacuation using three agent-based exit choice models. *Case studies on transport policy*, 9(3):1261–1272, 2021.

[78] Peter M Kielar and André Borrmann. Modeling pedestrians' interest in locations: A concept to improve simulations of pedestrian destination choice. *Simulation Modelling Practice and Theory*, 61:47–62, 2016.

[79] Wei-Li Wang, Fang-Fang Wan, and Siu-Ming Lo. Game theory model of exit selection in pedestrian evacuation considering visual range and choice firmness. *Chinese Physics B*, 29(8):084502, 2020.

[80] Miguel A Lopez-Carmona and Alvaro Paricio Garcia. Adaptive cell-based evacuation systems for leader-follower crowd evacuation. *Transportation research part C: emerging technologies*, 140:103699, 2022.

[81] Jur van den Berg, Sachin Patil, Dinesh Manocha, and Ming Lin. Interactive navigation of individual agents in crowded environments. *Interactive 3D Graphics and Games (I3D)*, 01 2008.

[82] BK Patle, Anish Pandey, DRK Parhi, AJDT Jagadeesh, et al. A review: On path planning strategies for navigation of mobile robot. *Defence Technology*, 15(4):582–606, 2019.

[83] Russell Gayle, Avneesh Sud, Ming C Lin, and Dinesh Manocha. Reactive deformation roadmaps: motion planning of multiple robots in dynamic environments. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3777–3783. IEEE, 2007.

[84] Jean-Claude Latombe. *Robot motion planning*, volume 124. Springer Science & Business Media, 2012.

[85] Ramon Oliva and Nuria Pelechano. Automatic generation of suboptimal navmeshes. In *International Conference on Motion in Games*, pages 328–339. Springer, 2011.

[86] Peng Zhang, Xiao-Xia Jian, SC Wong, and Keechoo Choi. Potential field cellular automata model for pedestrian flow. *Physical Review E*, 85(2):021119, 2012.

[87] Angel Garcimartín, Iker Zuriguel, Martin Pastor, César Martín-Gómez, and Daniel Parisi. Experimental evidence of the "faster is slower" effect. *Transportation Research Procedia*, 6:760–767, 10 2014.

[88] Peter Henry, Christian Vollmer, Brian Ferris, and Dieter Fox. Learning to navigate through crowded environments. In *2010 IEEE International Conference on Robotics and Automation*, pages 981–986. IEEE, 2010.

[89] Pete Trautman, Jeremy Ma, Richard M Murray, and Andreas Krause. Robot navigation in dense human crowds: Statistical models and experimental studies of human–robot cooperation. *The International Journal of Robotics Research*, 34(3):335–356, 2015.

[90] Chao Cao, Peter Trautman, and Soshi Iba. Dynamic channel: A planning framework for crowd navigation. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 5551–5557. IEEE, 2019.

[91] Stephen J Guy, Jatin Chhugani, Sean Curtis, Pradeep Dubey, Ming C Lin, and Dinesh Manocha. Pledestrians: A least-effort approach to crowd simulation. In *Symposium on computer animation*, pages 119–128, 2010.

[92] Sybren A. Stüvel, Nadia Magnenat-Thalmann, Daniel Thalmann, A. Frank van der Stappen, and Arjan Egges. Torso crowds. *IEEE Transactions on Visualization and Computer Graphics*, 23(7):1823–1837, 2017.

[93] Ignacio Mariano Sticco, Guillermo Alberto Frank, and Claudio Oscar Dorso. Social force model parameter testing and optimization using a high stress real-life situation. *Physica A: Statistical Mechanics and its Applications*, 561:125299, 2021.

[94] Mohamed Dridi. Simulation of high density pedestrian flow: A microscopic model. *Open Journal of Modelling and Simulation (OJMSi)*, 3:81–95, 08 2015.

[95] Ioannis Karamouzas, Nick Sohre, Rahul Narain, and Stephen J Guy. Implicit crowds: Optimization integrator for robust crowd simulation. *ACM Transactions on Graphics (TOG)*, 36(4):1–13, 2017.

[96] Tengfei Xu, Dongdong Shi, Juan Chen, Tao Li, Peng Lin, and Jian Ma. Dynamics of emotional contagion in dense pedestrian crowds. *Physics Letters A*, 384(3):126080, 2020.

[97] Gayathri Harihara Subramanian, Nipun Choubey, and Ashish Verma. Modelling and simulating serpentine group behaviour in crowds using modified social force model. *Physica A: Statistical Mechanics and its Applications*, 604:127674, 2022.

[98] Yushan Song, Xiangmin Hu, Liangchang Shen, and Wenguo Weng. Modeling domino effect along the queue using an improved social force model. *Physica A: Statistical Mechanics and its Applications*, 625:129008, 2023.

[99] Taras I Lakoba, David J Kaup, and Neal M Finkelstein. Modifications of the helbing-molnar-farkas-vicsek social force model for pedestrian evolution. *Simulation*, 81(5):339–352, 2005.

[100] Moonsoo Ko, Taewan Kim, and Keemin Sohn. Calibrating a social-force-based pedestrian walking model based on maximum likelihood estimation. *Transportation*, 40(1):91–107, 2013.

[101] Milad Haghani and Majid Sarvi. Simulating pedestrian flow through narrow exits. *Physics Letters A*, 383(2-3):110–120, 2019.

[102] Mohammed Mahmod Shuaib. Preserving socially expected crowd density in front of an exit for the reproduction of experimental data by modeling pedestrians' rear perception. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(10):P10037, 2014.

[103] Sahil Narang, Andrew Best, Sean Curtis, and Dinesh Manocha. Generating pedestrian trajectories consistent with the fundamental diagram based on physiological and psychological factors. *PLOS ONE*, 10:e0117856, 04 2015.

[104] Vitaly M Predtechenskii and Anatolii Ivanovich Milinskii. *Planning for foot traffic flow in buildings*. Amerind Publishing, New Dehli,, 1978.

[105] Masamitsu Mōri and Hiroshi Tsukaguchi. A new method for evaluation of level of service in pedestrian facilities. *Transportation Research Part A: General*, 21(3):223–234, 1987.

[106] Armin Seyfried, Bernhard Steffen, W. Klingsch, and Maik Boltes. The fundamental diagram of pedestrian movement revisited. *Journal of Statistical Mechanics: Theory and Experiment*, 10, 06 2005.

[107] Jun Zhang, W. Klingsch, A Schadschneider, and A Seyfried. Transitions in pedestrian fundamental diagrams of straight corridors and t-junctions. *Journal of Statistical Mechanics: Theory and Experiment*, 2011:P06004, 06 2011.

[108] Zheng Fang, S. Lo, and Jun-an Lu. On the relationship between crowd density and movement velocity. *Fire Safety Journal - FIRE SAFETY J*, 38:271–283, 04 2003.

[109] Ulrich Weidmann. Transporttechnik der fußgänger: transporttechnische eigenschaften des fußgängerverkehrs, literaturauswertung. *IVT Schriftenreihe*, 90, 1993.

[110] Ignacio Mariano Sticco, Guillermo Alberto Frank, Fernando Ezequiel Cornes, and Claudio Oscar Dorso. A re-examination of the role of friction in the original social force model. *Safety science*, 121:42–53, 2020.

[111] Wei-Li Wang, Hai-Cheng Li, Jia-Yu Rong, Qin-Qin Fan, Xin Han, and Bei-Hua Cong. A modified heuristics-based model for simulating realistic pedestrian movement behavior. *Chinese Physics B*, 31(9):094501, 2022.

[112] Fernando Alonso-Marroquin, Jonathan Busch, Coraline Chiew, Celia Lozano, and Álvaro Ramírez-Gómez. Simulation of counterflow pedestrian dynamics using spheropolygons. *Physical Review E*, 90(6):063305, 2014.

[113] Jingni Song, Feng Chen, Yadi Zhu, Na Zhang, Weiyu Liu, and Kai Du. Experiment calibrated simulation modeling of crowding forces in high density crowd. *IEEE Access*, 7:100162–100173, 07 2019.

[114] Francesco Zanlungo, Claudio Feliciani, Zeynep Yücel, Katsuhiro Nishinari, and Takayuki Kanda. Macroscopic and microscopic dynamics of a pedestrian cross-flow: Part ii, modelling. *Safety science*, 158:105969, 2023.

[115] Paolo Fiorini and Zvi Shiller. Motion planning in dynamic environments using the relative velocity paradigm. In *[1993] Proceedings IEEE International Conference on Robotics and Automation*, pages 560–565. IEEE, 1993.

[116] Paolo Fiorini and Zvi Shiller. Motion planning in dynamic environments using velocity obstacles. *The International Journal of Robotics Research*, 17:760–, 07 1998.

[117] Stephen Guy, Jatin Chhugani, Changkyu Kim, Nadathur Satish, Ming Lin, Dinesh Manocha, and Pradeep Dubey. Clearpath: Highly parallel collision avoidance for multi-agent simulation. In *Eurographics/ ACM SIGGRAPH Symposium on Computer Animation*, pages 177–187, 01 2009.

[118] Jamie Snape, Jur Van Den Berg, Stephen J Guy, and Dinesh Manocha. The hybrid reciprocal velocity obstacle. *IEEE Transactions on Robotics*, 27(4):696–706, 2011.

[119] Andrew Best, Sahil Narang, Sean Curtis, and Dinesh Manocha. Densesense: Interactive crowd simulation using density-dependent filters. In *Symposium on Computer Animation*, pages 97–102, 2014.

[120] Sean Curtis and Dinesh Manocha. *Pedestrian Simulation Using Geometric Reasoning in Velocity Space*, pages 875–890. Springer International Publishing, 2012.

[121] Sujeong Kim, Stephen Guy, and Dinesh Manocha. Velocity-based modeling of physical interactions in multi-agent simulations. *Proceedings - SCA 2013: 12th ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, 07 2013.

[122] Sujeong Kim, Stephen J. Guy, Karl Hillesland, Basim Zafar, Adnan Abdul-Aziz Gutub, and Dinesh Manocha. Velocity-based modeling of physical interactions in dense crowds. *Vis. Comput.*, 31(5):541–555, 5 2015.

[123] Sean Curtis, Stephen Guy, Basim Zafar, and Dinesh Manocha. Virtual tawaf: A case study in simulating the behavior of dense, heterogeneous crowds. In *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*, pages 128–135, 11 2011.

[124] Panich Sudkhot, K Wong, and Chattrakul Sombattheera. Collision avoidance and path planning in crowd simulation. *ICIC Express Lett*, 17:13, 2023.

[125] Anand S Rao, Michael P Georgeff, et al. Bdi agents: from theory to practice. In *Icmas*, volume 95, pages 312–319, 1995.

[126] Dhirendra Singh, Lin Padgham, and Brian Logan. Integrating bdi agents with agent-based simulation platforms. *Autonomous Agents and Multi-Agent Systems*, 30:1050–1071, 2016.

[127] Abhinav Golas, Rahul Narain, Sean Curtis, and Ming Lin. Hybrid long-range collision avoidance for crowd simulation. *IEEE transactions on visualization and computer graphics*, 20, 09 2013.

[128] Jingwei Ji, Ligang Lu, Zihao Jin, Shoupeng Wei, and Lu Ni. A cellular automata model for high-density crowd evacuation using triangle grids. *Physica A: Statistical Mechanics and its Applications*, 2018.

[129] Stefania Bandini, Matteo Mondini, and Giuseppe Vizzari. Modelling negative interactions among pedestrians in high density situations. *Transportation Research Part C-emerging Technologies*, 40:251–270, 2014.

[130] Ren-Yong Guo and Hai-Jun Huang. A modified floor field cellular automata model for pedestrian evacuation simulation. *Journal of Physics A: Mathematical and Theoretical*, 41(38):385104, 2008.

[131] Zhijian Fu, Qihao Jia, Junmin Chen, Jian Ma, Ke Han, and Lin Luo. A fine discrete field cellular automaton for pedestrian dynamics integrating pedestrian heterogeneity, anisotropy, and time-dependent characteristics. *Transportation research part C: emerging technologies*, 91:37–61, 2018.

[132] Grzegorz Bazior, Dariusz Pałka, and Jarosław Wąs. Using cellular automata to model high density pedestrian dynamics. In *International Conference on Computational Science*, pages 486–498. Springer International Publishing, 2020.

[133] Siamak Sarmady, Fazilah Haron, and Abdullah Zawawi Talib. Simulation of pedestrian movements using a fine grid cellular automata model. *IAES International Journal of Artificial Intelligence*, 11(4):1197, 2022.

[134] Claudio Feliciani and Katsuhiro Nishinari. An improved cellular automata model to simulate the behavior of high density crowd and validation by experimental data. *Physica A: Statistical Mechanics and its Applications*, 451, 02 2016.

[135] Feizhou Huo, Yufei Li, Chao Li, and Yaping Ma. An extended model describing pedestrian evacuation considering pedestrian crowding and stampede behavior. *Physica A: Statistical Mechanics and its Applications*, 604:127907, 2022.

[136] Qi Zhang, Jing Qu, and Yanzhe Han. Pedestrian small group behaviour and evacuation dynamics on metro station platform. *Journal of Rail Transport Planning & Management*, 26:100387, 2023.

[137] Ansgar Kirchner, Katsuhiro Nishinari, and Andreas Schadschneider. Friction effects and clogging in a cellular automaton model for pedestrian dynamics. *Physical review E*, 67(5):056122, 2003.

[138] Lili Lu, Ching-Yao Chan, Jian Wang, and Wei Wang. A study of pedestrian group behaviors in crowd evacuation based on an extended floor field cellular automaton model. *Transportation research part C: emerging technologies*, 81:317–329, 2017.

[139] Colin M Henein and Tony White. Agent-based modelling of forces in crowds. In *International Workshop on Multi-Agent Systems and Agent-Based Simulation*, pages 173–184. Springer, 2004.

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

[141] Jia-Bei Zeng, Biao Leng, Zhang Xiong, and Zheng Qin. Pedestrian dynamics in a two-dimensional complex scenario using a local view floor field model. *International Journal of Modern Physics C*, 22(08):775–803, 2011.

[142] Yushi Suma, Daichi Yanagisawa, and Katsuhiro Nishinari. Anticipation effect in pedestrian dynamics: Modeling and experiments. *Physica A: Statistical Mechanics and its Applications*, 391(1-2):248–263, 2012.

[143] Jakub Porzycki and Jarosław Wąs. Modeling spatial patterns in a moving crowd of people using data-driven approach—a concept of interplay floor field. *Safety science*, 167:106266, 2023.

[144] Ying Zheng, Xin-Gang Li, Bin Jia, and Rui Jiang. Simulation of pedestrians' evacuation dynamics with underground flood spreading based on cellular automaton. *Simulation Modelling Practice and Theory*, 94:149–161, 2019.

[145] Craig W Reynolds. Flocks, herds and schools: A distributed behavioral model. In *Proceedings of the 14th annual conference on Computer graphics and interactive techniques*, pages 25–34, 1987.

[146] Dirk Helbing. Agent-based modeling. In *Social self-organization*, pages 25–70. Springer, 2012.

[147] Min Zhou, Hairong Dong, Ding Wen, Xiuming Yao, and Xubin Sun. Modeling of crowd evacuation with assailants via a fuzzy logic approach. *IEEE Transactions on Intelligent Transportation Systems*, 17(9):2395–2407, 2016.

[148] Sharad Sharma, Kola Ogunlana, David Scribner, and Jock Grynovicki. Modeling human behavior during emergency evacuation using intelligent agents: A multi-agent simulation approach. *Information Systems Frontiers*, 20(4):741–757, 2018.

[149] Th Robin, Gianluca Antonini, Michel Bierlaire, and Javier Cruz. Specification, estimation and validation of a pedestrian walking behavior model. *Transportation Research Part B: Methodological*, 43(1):36–56, 2009.

[150] Jan Ondřej, Julien Pettré, Anne-Hélène Olivier, and Stéphane Donikian. A synthetic-vision based steering approach for crowd simulation. *ACM Transactions on Graphics (TOG)*, 29(4):1–9, 2010.

[151] Julio Erasmo Godoy, Ioannis Karamouzas, Stephen J Guy, and Maria Gini. Implicit coordination in crowded multi-agent navigation. In *Thirtieth AAAI Conference on Artificial Intelligence*, pages 2487–2493, 2016.

[152] Liang He, Jia Pan, Wenping Wang, and Dinesh Manocha. Proxemic group behaviors using reciprocal multi-agent navigation. In *2016 IEEE international conference on robotics and automation (ICRA)*, pages 292–297. IEEE, 2016.

[153] Hong Liu, Baoxi Liu, Hao Zhang, Liang Li, Xin Qin, and Guijuan Zhang. Crowd evacuation simulation approach based on navigation knowledge and two-layer control mechanism. *Information Sciences*, 436:247–267, 2018.

[154] Maosheng Li, Panpan Shu, Yao Xiao, and Pu Wang. Modeling detour decision combined the tactical and operational layer based on perceived density. *Physica A: Statistical Mechanics and its Applications*, 574:126021, 2021.

[155] Wei Xie, Eric Wai Ming Lee, and Yiu Yin Lee. Simulation of spontaneous leader–follower behaviour in crowd evacuation. *Automation in Construction*, 134:104100, 2022.

[156] Alex Graves and Alex Graves. Long short-term memory. *Supervised sequence labelling with recurrent neural networks*, pages 37–45, 2012.

[157] Namhoon Lee, Wongun Choi, Paul Vernaza, Christopher B Choy, Philip HS Torr, and Manmohan Chandraker. Desire: Distant future prediction in dynamic scenes with interacting agents. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 336–345, 2017.

[158] Anirudh Vemula, Katharina Muelling, and Jean Oh. Social attention: Modeling attention in human crowds. In *2018 IEEE international Conference on Robotics and Automation (ICRA)*, pages 4601–4607. IEEE, 2018.

[159] Amir Sadeghian, Vineet Kosaraju, Ali Sadeghian, Noriaki Hirose, Hamid Rezatofighi, and Silvio Savarese. Sophie: An attentive gan for predicting paths compliant to social and physical constraints. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1349–1358, 2019.

[160] Kai Zhu, Bin Li, Wenming Zhe, and Tao Zhang. Collision avoidance among dense heterogeneous agents using deep reinforcement learning. *IEEE Robotics and Automation Letters*, 8(1):57–64, 2022.

[161] Parth Kothari, Sven Kreiss, and Alexandre Alahi. Human trajectory forecasting in crowds: A deep learning perspective. *IEEE Transactions on Intelligent Transportation Systems*, 23(7):7386–7400, 2021.

[162] Bengt Andersson, Ronnie Andersson, Love Håkansson, Mikael Mortensen, Rahman Sudiyo, and Berend Van Wachem. *Computational fluid dynamics for engineers*. Cambridge university press, 2011.

[163] Yan-qun Jiang, Peng Zhang, SC Wong, and Ru-xun Liu. A higher-order macroscopic model for pedestrian flows. *Physica A: Statistical Mechanics and its Applications*, 389(21):4623–4635, 2010.

[164] Yan-Qun Jiang, Shu-Guang Zhou, and Fang-Bao Tian. A higher-order macroscopic model for bi-direction pedestrian flow. *Physica A: Statistical Mechanics and its Applications*, 425:69–78, 2015.

[165] Roger L Hughes. The flow of human crowds. *Annual review of fluid mechanics*, 35(1):169–182, 2003.

[166] Wouter van Toll, Cédric Braga, Barbara Solenthaler, and Julien Pettré. Extreme-density crowd simulation: Combining agents with smoothed particle hydrodynamics. In *Proceedings of the 13th ACM SIGGRAPH Conference on Motion, Interaction and Games*, pages 1–10, 2020.

[167] Yufei Yuan, Bernat Goñi-Ros, Ha H Bui, Winnie Daamen, Hai L Vu, and Serge P Hoogendoorn. Macroscopic pedestrian flow simulation using smoothed particle hydrodynamics (sph). *Transportation research part C: emerging technologies*, 111:334–351, 2020.

[168] Wouter van Toll, Thomas Chatagnon, Cédric Braga, Barbara Solenthaler, and Julien Pettré. Sph crowds: Agent-based crowd simulation up to extreme densities using fluid dynamics. *Computers & Graphics*, 98:306–321, 2021.

[169] Robert A Gingold and Joseph J Monaghan. Smoothed particle hydrodynamics: theory and application to non-spherical stars. *Monthly notices of the royal astronomical society*, 181(3):375–389, 1977.

[170] A Jebrane, P Argoul, A Hakim, and M El Rhabi. Estimating contact forces and pressure in a dense crowd: Microscopic and macroscopic models. *Applied Mathematical Modelling*, 74:409–421, 2019.

[171] Qiang Chen, Guoliang Luo, Yang Tong, Xiaogang Jin, and Zhigang Deng. A linear wave propagation-based simulation model for dense and polarized crowds. *Computer Animation and Virtual Worlds*, 32(1):e1977, 2021.

[172] Jibiao Zhou, Siyuan Chen, Changxi Ma, and Sheng Dong. Stability analysis of pedestrian traffic flow in horizontal channels: A numerical simulation method. *Physica A: Statistical Mechanics and its Applications*, 587:126528, 2022.

[173] Rongyong Zhao, Qianshan Hu, Qiong Liu, Cuiling Li, Daheng Dong, and Yunlong Ma. Panic propagation dynamics of high-density crowd based on information entropy and aw-rascle model. *IEEE Transactions on Intelligent Transportation Systems*, 21(10):4425–4434, 2019.

[174] Haoyang Liang, Jie Du, and Sze Chun Wong. A continuum model for pedestrian flow with explicit consideration of crowd force and panic effects. *Transportation research part B: methodological*, 149:100–117, 2021.

[175] Cuiling Li, Rongyong Zhao, Chuanfeng Han, Rahman Arifur, Yunlong Ma, and Qiong Liu. Dynamic disturbance propagation model of pedestrian panic behaviors and lyapunov-based crowd stability analysis. *Applied Sciences*, 13(21):11762, 2023.

[176] Wenjie Zhu, Rongyong Zhao, Hao Zhang, Cuiling Li, Ping Jia, Yunlong Ma, Dong Wang, and Miyuan Li. Panic-pressure conversion model from microscopic pedestrian movement to macroscopic crowd flow. *Journal of Computational and Nonlinear Dynamics*, 18(12), 2023.

[177] Nur Siyam, Omar Alqaryouti, and Sherief Abdallah. Research issues in agent-based simulation for pedestrians evacuation. *IEEE Access*, 8:134435–134455, 2019.

[178] Muzhou Xiong, Shanyu Tang, and Dan Zhao. A hybrid model for simulating crowd evacuation. *New Generation Computing*, 31:211–235, 2013.

[179] Chighoub Rabiaa and Cherif Foudil. Toward a hybrid approach for crowd simulation. *International Journal of Advanced Computer Science and Applications*, 7(1), 2016.

[180] Daniel H Biedermann, Jan Clever, and Andre Borrmann. A generic and density-sensitive method for multi-scale pedestrian dynamics. *Automation in Construction*, 122:103489, 2021.

[181] Muzhou Xiong, Wentong Cai, Suiping Zhou, Malcolme Yoke Hean Low, Feng Tian, Dan Chen, Darren Wee Sze Ong, and Benjamin D Hamilton. A case study of multi-resolution modeling for crowd simulation. In *Proceedings of the 2009 Spring Simulation Multiconference*, pages 1–8, 2009.

[182] Simone Göttlich and Marion Pfirsching. A micro-macro hybrid model with application for material and pedestrian flow. *Cogent Mathematics & Statistics*, 5(1):1476049, 2018.

[183] Hoshang Kolivand, Mohd Shafry Rahim, Mohd Shahrizal Sunar, Ahmad Zakwan Azizul Fata, and Chris Wren. An integration of enhanced social force and crowd control models for high-density crowd simulation. *Neural Computing and Applications*, 33:6095–6117, 2021.

[184] Sahil Narang, Andrew Best, and Dinesh Manocha. Interactive simulation of local interactions in dense crowds using elliptical agents. *Journal of Statistical Mechanics: Theory and Experiment*, 2017:033403, 03 2017.

[185] Jian Ma, Wei-guo Song, Jun Zhang, Siu-ming Lo, and Guang-xuan Liao. k-nearest-neighbor interaction induced self-organized pedestrian counter flow. *Physica A: Statistical Mechanics and its Applications*, 389(10):2101–2117, 2010.

[186] Xinwei Zhang, Peihong Zhang, and Maohua Zhong. A dual adaptive cellular automaton model based on a composite field and pedestrian heterogeneity. *Physica A: Statistical Mechanics and its Applications*, 583:126334, 2021.

[187] Yan-Qun Jiang, Ying-Gang Hu, and Xiaoqian Huang. Modeling pedestrian flow through a bottleneck based on a second-order continuum model. *Physica A: Statistical Mechanics and its Applications*, 608:128272, 2022.

[188] Stefano Pellegrini, Andreas Ess, Konrad Schindler, and Luc Van Gool. You'll never walk alone: Modeling social behavior for multi-target tracking. In *2009 IEEE 12th international conference on computer vision*, pages 261–268. IEEE, 2009.

[189] Alon Lerner, Yiorgos Chrysanthou, and Dani Lischinski. Crowds by example. In *Computer graphics forum*, pages 655–664. Wiley Online Library, 2007.

[190] Shuchao Cao, Armin Seyfried, Jun Zhang, Stefan Holl, and Weiguo Song. Fundamental diagrams for multidirectional pedestrian flows. *Journal of Statistical Mechanics: Theory and Experiment*, 2017(3):033404, 2017.

[191] Festival of lights. `https://www.fetedeslumieres.lyon.fr`. Accessed on 31 May 2024.

[192] Data archive of studies about pedestrian dynamics. `https://ped.fz-juelich.de/da/doku.php?id=start`. Accessed on 24 September 2024.

[193] Maik Boltes, Ann Katrin Boomers, Juliane Adrian, Ricardo Martin Brualla, Arne Graf, Paul Häger, Daniel Hillebrand, Deniz Kilic, Paul Lieberenz, Daniel Salden, and Tobias Schrödter. Petrack, July 2021.

[194] Maik Boltes and Armin Seyfried. Collecting pedestrian trajectories. *Neurocomputing*, 100:127–133, 2013.

[195] Bruce D Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI'81: 7th international joint conference on Artificial intelligence*, volume 2, pages 674–679, 1981.

[196] Zhengyou Zhang. A flexible new technique for camera calibration. *IEEE Transactions on pattern analysis and machine intelligence*, 22(11):1330–1334, 2000.

[197] Christopher M Bishop and Nasser M Nasrabadi. *Pattern recognition and machine learning*, volume 4. Springer, 2006.

[198] Mohcine Chraibi. madras-data-app, February 2024.

[199] Dirk Helbing. A fluid dynamic model for the movement of pedestrians. *arXiv preprint cond-mat/9805213*, 1998.

[200] Eric Freeman, Elisabeth Robson, Bert Bates, and Kathy Sierra. *Head First Design Patterns: A Brain-Friendly Guide*. " O'Reilly Media, Inc.", 2004.

[201] Anders Johansson, Dirk Helbing, and Pradyumn K Shukla. Specification of the social force pedestrian model by evolutionary adjustment to video tracking data. *Advances in complex systems*, 10(supp02):271–288, 2007.

[202] Mehdi Moussaïd, Dirk Helbing, Simon Garnier, Anders Johansson, Maud Combe, and Guy Theraulaz. Experimental study of the behavioural mechanisms underlying self-organization in human crowds. *Proceedings of the Royal Society B: Biological Sciences*, 276(1668):2755–2762, 2009.

[203] Patrick Taillandier, Benoit Gaudou, Arnaud Grignard, Quang-Nghi Huynh, Nicolas Marilleau, Philippe Caillou, Damien Philippon, and Alexis Drogoul. Building, composing and experimenting complex spatial models with the gama platform. *GeoInformatica*, 23:299–322, 2019.

[204] Serge P Hoogendoorn, Winnie Daamen, and R Landman. Microscopic calibration and validation of pedestrian models—cross-comparison of models using experimental data. In *Pedestrian and Evacuation Dynamics 2005*, pages 253–265. Springer, 2007.

[205] Marion Gödel, Nikolai Bode, Gerta Köster, and Hans-Joachim Bungartz. Bayesian inference methods to calibrate crowd dynamics models for safety applications. *Safety science*, 147:105586, 2022.

[206] Weiliang Zeng, Peng Chen, Guizhen Yu, and Yunpeng Wang. Specification and calibration of a microscopic model for pedestrian dynamic simulation at signalized intersections: A hybrid approach. *Transportation Research Part C: Emerging Technologies*, 80:37–70, 2017.

[207] Xu Chen, Martin Treiber, Venkatesan Kanagaraj, and Haiying Li. Social force models for pedestrian traffic–state of the art. *Transport reviews*, 38(5):625–653, 2018.

[208] Dirk Helbing, Anders Johansson, and Habib Zein Al-Abideen. Dynamics of crowd disasters: An empirical study. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 75(4):046109, 2007.

[209] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A density-based algorithm for discovering clusters in large spatial databases with noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, page 226–231. AAAI Press, 1996.

[210] Hao Xue, Du Q Huynh, and Mark Reynolds. Ss-lstm: A hierarchical lstm model for pedestrian trajectory prediction. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1186–1194. IEEE, 2018.

[211] Sirin Haddad, Meiqing Wu, He Wei, and Siew Kei Lam. Situation-aware pedestrian trajectory prediction with spatio-temporal attention model. *24th Computer Vision Winter Workshop*, pages 4–16, 2019.

[212] Yingfan Huang, Huikun Bi, Zhaoxin Li, Tianlu Mao, and Zhaoqi Wang. Stgat: Modeling spatial-temporal interactions for human trajectory prediction. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 6272–6281, 2019.

[213] Alessio Monti, Alessia Bertugli, Simone Calderara, and Rita Cucchiara. Dag-net: Double attentive graph neural network for trajectory forecasting. In *2020 25th International Conference on Pattern Recognition (ICPR)*, pages 2551–2558. IEEE, 2021.

[214] Wei-Cheng Lai, Zi-Xiang Xia, Hao-Siang Lin, Lien-Feng Hsu, Hong-Han Shuai, I-Hong Jhuo, and Wen-Huang Cheng. Trajectory prediction in heterogeneous environment via attended ecology embedding. In *Proceedings of the 28th acm international conference on multimedia*, pages 202–210, 2020.

[215] Ioannis Karamouzas, Brian Skinner, and Stephen J Guy. Universal power law governing pedestrian interactions. *Physical review letters*, 113(23):238701, 2014.

[216] S Hochreiter. Long short-term memory. *Neural Computation MIT-Press*, 1997.

[217] Armin Seyfried, Oliver Passon, Bernhard Steffen, Maik Boltes, Tobias Rupprecht, and Wolfram Klingsch. New insights into pedestrian flow through bottlenecks. *Transportation Science*, 43(3):395–406, 2009.

[218] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences*, 108(17):6884–6888, 2011.

[219] Antoine Tordeux, Mohcine Chraibi, Armin Seyfried, and Andreas Schadschneider. Prediction of pedestrian dynamics in complex architectures with artificial neural networks. *Journal of intelligent transportation systems*, 24(6):556–568, 2020.

List of Figures

2.1 Flow diagram for collecting articles relevant to the modeling of pedestrian behavior in high-density situations. 17
2.2 Analysis results on the collected articles. 19
2.3 Number of publications per year for dense crowd simulations. 19
2.4 Word cloud extracted from the abstract and authors' keywords of collected articles. 20
2.5 Variations in the number of studies over the two periods: 2000 – 2011 and 2012 – 2023. 21
2.6 The classical Perception-Decision-Action schema. 21
2.7 Different tactical-level models for pedestrian navigation: (a) Graph network generated based on Voronoi edges (white lines); (b) Navigation mesh constructed using the medial axis (blue lines) and orange lines connecting the medial axis vertices to proximal obstacles; (c) Navigation field describing preferred moving directions of crowds across different local areas. 24
2.8 Hierarchical classification of operational-level models. 27
2.9 Velocity obstacle models [116, 62, 63]. 31
2.10 Combination of FSM and RVO in [123]. 33
2.11 Finer representations for pedestrian shape. 34
2.12 Multi-scale simulation. 55

3.1 Festival of Lights in 2022. 59
3.2 Experimental design. 61
3.3 Screenshots of all cameras. 62
3.4 Image from camera 3 including reference points for extrinsic calibration. 64

3.5 Region of interest defined by a yellow rectangle for videos recorded by camera 3. 65
3.6 Reference lines to count pedestrian heads. 66
3.7 TopView_2D 68
3.8 Time-series outflow of pedestrians on Chenavard road and Constantine road. . . 69
3.9 Similar regions of interest in the TopView_1B and TopView_2C camera videos, which capture the same scene but from different locations and angles. 70
3.10 Partial overlap of the average speed and density sequences over time for two pairs of video recordings: TopView_1B/TopView_2C and TopView_1C/TopView_-2D, which capture the same scene from different positions and perspectives. . . 71

4.1 General overview. 77
4.2 Architecture of the HyPedSim framework. 79
4.3 Pedestrian activity diagram at each simulation step. 81
4.4 1 82
4.5 Description of repulsion force f_{ij} between pedestrian i and j. 84
4.6 Parameters of CC model. 86
4.7 Discretization of the environment into three zones. 86
4.8 Illustration of uniform crossover process between two parental chromosomes. . 91
4.9 Flow of the calibration process. 92
4.10 Average and best fitness values over 150 generations. 93
4.11 Comparison of observed and simulated outflow for the two exit roads. (**a**) Constantine road. (**b**) Chenavard road. 94
4.12 *Cont.* 96
4.12 Local sensitivity analysis for different parameters. (**a**) t_{delay}, (**b**) α, (**c**) β, (**d**) A, (**e**) B, (**f**) V^{pref}, (**g**) τ, (**h**) $f_{\min}$, (**i**) $f_{\max}$, (**j**) $\rho_{\min}$, (**k**) $\rho_{\max}$. 97
4.13 Density maps among different models with 6000 simulated agents. 99
4.14 A snapshot of simulation of the hybrid model with 6000 agents. 100
4.15 Comparison of the performance of different models. 101

5.1 Geometric illustration of TTC calculation. 108
5.2 Example of computing TTC. 109
5.3 Interaction energy with $k = 1.5$ and $\tau_0 = 3.0$. 111
5.4 SLSTM architecture developed in [11]. 112
5.5 Example illustrating limitation in predictions. 113
5.6 $\tanh$ function of the interaction energy $E(\tau)$ with $k = 1.5$ and $\tau_0 = 3.0$. 114

5.7 Trajectory examples from the datasets. 116
5.8 Average evaluation results over different values of λ on the low-density datasets. 123
5.9 Average evaluation results over different values of λ on the high-density datasets. 124
5.10 Prediction examples of SLSTM and TTC-SLSTM. 125

List of Tables

2.1 Main classes of models at different levels of pedestrian behavior. 18

2.2 Comparison of the characteristics of tactical level models for simulating pedestrian tactical behavior in high-density scenarios (CT: Construction time, MUT: Map update time, PST: Path search time, NA: Number of agents, RT: Run time). 26

2.3 Comparison results of operational level models based on various evaluation criteria (PI: physical interaction, GB: group behavior, FB: following behavior, RB: replan behavior, TS: time step (s), MD: max density ($\mathrm{ped}/\mathrm{m}^2$), NA: number of agents, RT: run time (fps), S: strategic, T: tactical, O: operational). 45

2.4 Assessments of the advantages and disadvantages of operational-level models for simulating high-density crowds (PI: physical interaction, GB: group behavior, FB: following behavior, RB: replan behavior, TSA: time step adaptivity, D: density, SS: simulation speed, S: scalability, IHM: integration of high-level modeling, AHM: ability of high-level modeling). 50

3.1 File tracking, timestamps, trajectory counts, and main statistics for the TopView dataset. 67

3.2 Root mean square errors (RMSE) between the pedestrian mean speed and density time series for the TopView_1B/TopView_2C and TopView_1C/TopView_2D video recordings. 70

4.1 List of parameters for calibration and their ranges of values. 90

4.2 Best solution obtained through the calibration process using the GA method. . . 93

5.1 Fundamental characteristics of different datasets. 115

5.2 Prediction results of different models on low-density datasets. 121
5.3 Prediction results of different models on high-density datasets. 122