Modeling Unidirectional Pedestrian Movement: An Investigation of Diffusion Behavior in the Built Environment

Ying Liu,¹,² Cheng Sun,¹ and Yiming Bie³

¹School of Architecture, Harbin Institute of Technology, Harbin 150001, China
²Center for Green Buildings and Cities, Graduate School of Design, Harvard University, Cambridge, MA 02138, USA
³School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150091, China

Correspondence should be addressed to Yiming Bie; yimingbie@126.com

Received 3 June 2015; Revised 9 August 2015; Accepted 25 August 2015

1. Introduction

Pedestrian Movement Theories. Understanding pedestrian movement is essential for the safety planning and design of facilities. Pedestrian flow theories and models can support infrastructure designers and transport planners to optimize their plans and efficiently control pedestrian traffic and help architects facilitate circulation design in and around assembly occupancies, such as sports venues [1], transportation facilities and interchanges [2], and holy sites and political demonstrations [3]. To meet the increasing safety demands in organizing large scale public events, pedestrian flow theories are often introduced to improve the techniques and effectiveness of crowd management and control. In the design practice of pedestrian facilities, unidirectional movement such as the promotion of one-way traffic in urban road networks and separation of ingress and egress at airports, stations, subways, stadiums, exhibition centers, and supermarkets are considered to be beneficial for efficiently increasing pedestrian flow while maintaining the circulation safety. Unimpeded unidirectional flow keeps pedestrian crowds moving in steady state, which is also the most desirable circumstance in the evacuation process of large crowds.

Pedestrian movement theories are mostly planning related, with large urban and neighborhood scale supporting experimentations. Early studies on pedestrian behavior and movement were mostly empirical, and regression analysis has been widely used for proposing theories, understanding and evaluating design elements of pedestrian facilities, and developing planning guidelines [4]. Fruin proposed the level-of-service (LoS) concept and standard based on a range of pedestrian area occupancies [5]. Pauls studied pedestrian movement in buildings and design solutions of means of egress in cases involving both traditional total evacuation and selective sequential evacuation of large office buildings during fire drills for understanding the performance of the exit geometry, communication, operation, and training [6, 7]. Over the past twenty years, methods for describing and quantifying pedestrian movement attracted much attention from scholars around the world. A number of simulation models have been proposed and widely used for the prediction of pedestrian flow in the planning and design process, such as...
as Simulex [8], Building Exodus [9], Legion [10], and Space Syntax [11]. Up until now, because of the inherent complexity of interpreting collective behavior, research efforts have continuously been evolving to improve the efficiency and accuracy of prediction models for pedestrian movement.

**State-of-the-Art Methods.** A number of computational methods focusing on the pedestrian movement have been introduced and simulation models have been developed in both academia and industry. Existing methods can be divided into two major categories: microscopic and macroscopic. Microscopic models simulate pedestrians as a set of individuals or agents and describe individualized behaviors and interactions in greater detail, which requires greater computational power. Macroscopic models, on the other hand, focus on overall pedestrian movement situations and view pedestrians as flows. These models can be used for minimum travel time estimation and significantly save computational time. In addition, there are some studies using the mesoscopic approach for pedestrian simulation in public buildings [12]. The idea is mainly based on grouping individuals. Physical groups can serve as input data, while virtual or logical groups are used to show similar behaviors or intentions [13].

The lattice-gas models, fluid-dynamic models, and network flow models commonly use macroscopic modeling approaches. Lattice gas models consider each pedestrian as an active particle on the grid for studying the features of pedestrian crowds by means of probability and statistics. Helbing et al. proposed a lattice gas model of pedestrian flows and compared the results with experimental data [14]. Fluid-dynamic models describe pedestrian crowds with fluid-like properties that behave similarly to gases or fluids. Helbing presented a model for pedestrian collective movement based on Boltzmann-like gas-kinetic approach [15]. Hughes presented a model based on three hypothesis from well-defined observations [16]. These fluid-dynamic models often describe density and velocity change over time using partial differential equations [17]. The network flow model can be also used to estimate pedestrian travel time and explore optimal routes. Hamacher et al. proposed a model which considered the quickest cluster flow problem based on network flow theories [18]. In addition, other common macroscopic models include the regression model, route choice model, queuing model, and gas-kinetics model.

Microscopic models can be divided into several categories: physical-based models, cellular automata models (CA), rule-based models, and multiagent models. Physical-based models recognize crowds as individuals reacting to surrounding events and have been used to study the emergency and panic situations, where optimal acceleration is controlled by varying physical force. The magnetic force model, NOMAD (Normative Pedestrian Behavior Theory), and social force model are examples of physical-based models [12]. CA models, also known as matrix-based models, divide a building floor into grids or cells. Each individual evacuee occupies one or multiple cells and moves forward when there are empty cells available. CA models can be further divided into either coarse-grid models or fine-grid models based on the size of the grids. Rule-based models have been widely used in computer graphics for modeling groups of animals and virtual crowds. An example of this is Reynolds’ “birds” model which followed three local rules [19]. These rules separation, alignment, and cohesion are used to simulate flocking behavior in the model. Multiagent models allow the environment to be populated with interactive and autonomous agents, which is a preferred approach for large scale and complex evacuation simulation.

**Current Barriers and Proposed Model.** Despite the numerous models that have been developed for simulating pedestrian movement, limitations still exist in interpreting collective behavior. Particularly in macroscopic methods, only a few studies integrate behavioral impact into the model. Hughes developed a continuum model to describe crowds as “thinking fluids” [16]. Colombo and Rosini proposed a model for pedestrian flow to describe typical features of flow under the effect of panic [20]. Helbing et al. investigated a model to simulate the dynamic feature of the disastrous form of collective human behavior caused by panic and jam [21]. The numerically efficient cellular automata (CA) models treat pedestrians as interacting particles. Some CA models allow partial behavior, which is a compromise between the macroscopic and microscopic approaches [22]. Diffusion phenomena have also been observed and specified in models for vehicle transportation [23, 24]. However, the existence of diffusion behavior in unidirectional pedestrian crowd movement has neither been noticed nor addressed yet in current models, which can be important in representing spatial-temporal crowd density differentiation in the movement process.

This paper proposes a new approach that incorporates diffusion behavior into modeling of pedestrian movement. The "spread-out" phenomenon that pedestrian crowds display during the unidirectional movement is illustrated in algorithms for pedestrian flow estimation. The proposed methodology features discrete-time, geometrical speed distribution and degree of diffusion in relation to speed divergence, which can be used to calculate pedestrian arrival state at any forward position and analyze unidirectional pedestrian walking processes from the macroscopic perspective. The model is then validated using video tracking data for the evacuation process of a long egress ramp in a stadium. The minimum difference value between the actual video sampling data and estimated results in various time intervals is used as an evaluation index to achieve the optimized degree of diffusion. The verification results show that the proposed model could accurately describe the pedestrian movement on the passage and makes the macroscopic pedestrian movement modeling approach more realistic by taking diffusion behavior into consideration.

**2. Methodology**

**2.1. Diffusion Model.** The diffusion modeling concept is illustrated in Figure 1. As shown in the figure, as a pedestrian crowd moves from left to right along a unidirectional long passage, both crowd density and flow rate are relatively higher at time $t_1$ segment $A$. However, during the movement
process towards the right direction, because of walking speed differentiations among individual pedestrians, some faster individuals in the group will overtake the slower ones. The interpersonal distances among each pedestrian will increase gradually, while pedestrian density per unit will decrease. Therefore, pedestrian flow rate at time \( t_2 \) \((t_2 > t_1)\) segment \( B \) will be smaller than segment \( A \). This observed "spread-out" phenomena of the walkers can be described as crowd diffusion. The crowd diffusion phenomenon commonly exists in buildings and urban streets. Studying crowd diffusion behavior can be beneficial for analyzing pedestrian flow rate and density in buildings or on streets and provides a method for evaluating the rationality of level of service (LoS) and pedestrian facilities.

2.2. Modeling Procedure. The proposed model partitions time into multiple intervals. The length of each time interval is \( \Delta t \) seconds. Time 0 is defined by the time when the first pedestrian in the group walks through segment \( A \), hence \( T_A \) represents the moment when the last pedestrian walk through segment \( A \), and the total number of time intervals is \( N_A \):

\[
N_A = \text{Ceiling} \left( \frac{T_A}{\Delta t} \right),
\]  

(1)

where Ceiling is defined as a top integral function. For example, both Ceiling (4.2) and Ceiling (4.7) are equal to 5.

Then, the number of pedestrians that walk through segment \( A \) is counted. Taking time interval \( i \) \((1 \leq i \leq N_A)\) as an example, the number of pedestrians walking through segment \( A \) is \( q_A(i) \). \( V \) (m/s) represents the average pedestrian walking speed. However, there are differences between the walking speeds of various pedestrians, assuming all pedestrian walking speeds follow geometric distribution.

For the \( i \) pedestrians in one interval, their speeds can be calculated according to

\[
V_i = V_{\max} F (1 - F)^{i-1},
\]  

(2)

where \( F \) is the smoothing coefficient and \( V_{\max} \) is the maximum speed. The determination method for \( V_{\max} \) will be illustrated in the following paragraphs.

When the pedestrians arrive at the downstream segment \( B \), the arrival flow rate should follow geometric distribution as well. The predicted number of pedestrians \( q_B(j) \) that will arrive at segment \( B \) at the \( j \)th time interval equals

\[
q_B'(j) = F \cdot q_A(j - T) + F(1 - F) \cdot q_A(j - 1 - T) + F(1 - F)^2 q_A(j - 2 - T) + \cdots
\]  

\[
+ F(1 - F)^k q_A(j - n - T) + \cdots
\]  

\[
+ F(1 - F)^{j-T-1} q_A(1),
\]  

(3)

where \( T \) represents the walking time of the pedestrian with fastest speed from segment \( A \) to segment \( B \), expressed by time intervals:

\[
F = \frac{1}{1 + \gamma_1 \cdot \gamma_2 \cdot \delta_a},
\]  

(4)

\[
T = \gamma_2 \cdot \delta_a,
\]  

(5)

where \( \gamma_1 \) is the diffusion coefficient; \( \gamma_2 \) is the travel time coefficient; and \( \delta_a \) is the average walking time for all pedestrians from segment \( A \) to segment \( B \), which equals \( L_{AB}/V \), \( L_{AB} \) is the distance between segment \( A \) and segment \( B \) in meters.

\( F \) is an important parameter in this model, as it can be used to depict the diffusion level of the pedestrians. It should be calibrated in practice by considering various contributing factors, such as the road width, road grade, pedestrian flow volume, and pedestrian types (e.g., male, female, teenagers, children, and the elderly). In this study, \( F \) is affected by \( \gamma_1 \) and \( \gamma_2 \). Different values of \( \gamma_1 \) and \( \gamma_2 \) will result in a different \( F \) value. In Section 3.2, a calibration method is brought forward to determine the optimal values for \( \gamma_1 \) and \( \gamma_2 \), which will also produce the optimal value for \( F \).

In (4) and (5), \( \gamma_1 \) represents the degree of pedestrian diffusion, which is a positive number less than 1. Equation (4) shows that, as \( \gamma_1 \) increases, the \( F \) value decreases, meaning that the larger the diffusion coefficient, the lower the degree of pedestrian diffusion.

\( \gamma_2 \) is a positive number smaller than 1. This parameter is introduced because there are differences in pedestrian walking speeds. Those who walk faster will arrive at segment \( B \) earlier than other pedestrians. Parameter \( T \) is the walking time of the fastest pedestrian from segment \( A \) to segment \( B \). In other words, the pedestrian with the largest speed at segment \( A \) during time interval \( i \) will arrive at segment \( B \) in the \( (i+T) \) time interval. In addition, the fastest walking speed is \((1/\gamma_2)\) times the average walking speed.

Thus, the maximum walking speed \( V_{\max} \) (m/s) equals

\[
V_{\max} = \frac{L_{AB}}{\gamma_2 \cdot \delta_a \cdot \Delta t},
\]  

(6)

In (6), \( \Delta t \) is the number of seconds in one unit interval and \( \gamma_2 \cdot \delta_a \cdot \Delta t \) equals the minimum travel time between segments \( A \) and \( B \).
To further develop the formula, (7) can be as follows:

\[ q_B^e(j) = F \cdot q_A(j-T) + (1-F) \cdot q_B^e(j-1). \]  

Equation (7) estimates the number of pedestrians at each time interval in the downstream segment B based on the number of pedestrians and the average pedestrian walking speed in each time interval from the upstream segment A. Segment B can be set up at any location downstream from segment A. Therefore, (6) can be used to calculate pedestrian arrival state at any location in front of segment A and to analyze the walking process of the pedestrian stream on the passage from macroperspectives.

### 2.3. Model Features

The proposed crowd movement diffusion model has the following features.

#### 2.3.1. Discretization of Time Variable

In the modeling process, time is discretized into multiple intervals, the unit length of which is defined in seconds. Because of the individual differences in gait parameters and walking speeds, diffusion phenomena exists in each time interval. Diffusion behavior of the whole pedestrian stream consists of crowd diffusion behavior in each time interval. When compared to the continuous model, the discrete model is more easily translated into programming and the computation resource requirement is relatively low. This allows the model to satisfy the requirements of large scale software development.

#### 2.3.2. Walking Speed in Geometric Distribution

Pedestrian stream diffusion occurs because each pedestrian in the crowd has his/her own walking velocity. If all pedestrians walk at the same speed, then there will be no diffusion behavior. Therefore, the greater the individual pedestrian velocity differences are, the more serious the pedestrian stream diffusion can be, which means that as the value of \( \gamma_1 \) increases, the value of \( F \) decreases accordingly. This study describes the pedestrian crowd walking speed distribution by a geometrical distribution model. It cannot only describe the pedestrian walking speed differences, but also adjust the modeling structure by parameter \( F \) to make the geometrical model more practical and realistic.

#### 2.3.3. Key Parameter Values and Impact

The degree of diffusion and walking time of the fastest person in each time interval are the key parameters. In the equations, \( \gamma_1 \) represents the diffusion degree of the pedestrian stream, which is influenced by the divergence degree of each pedestrian's walking speed. \( \gamma_2 \) equals the walking time of the pedestrian with the fastest speed in each time interval from the upstream segment A to downstream segment B, represented by the number of time intervals. \( \gamma_2 \) is a positive number less than or equal to 1.0. The higher the value is, the closer the fastest walking speed is to the average pedestrian walking speed; the lower the value is, the greater the fastest walking speed is compared to the average pedestrian walking speed. If \( \gamma_2 \) is equal to 0.8, this indicates that the greater fastest walking speed is 1.25 (1.0/0.8) times larger than the average walking speed \( V \).

### 3. Experimentation

#### 3.1. Data Collection

To verify the validity of the proposed model, we collected data of pedestrian movements on a field passage to verify the model precision. Since the values of parameters \( \gamma_1 \) and \( \gamma_2 \) have a large fluctuation range, this section will also verify the impact of these parameters on model precision.

Video recordings were conducted on the long egress ramp of Citizens Bank Park in Philadelphia (Pennsylvania, United States) during the spectators' evacuation process. Citizens Bank Park is a baseball stadium built in 2004, the capacity of which is 43.500. The playing field is 7 m lower than the ground level. Spectators from the upper stands could evacuate to the ground by a ramp, escalators, and elevators. The observed match was very popular. All the seats had been occupied since the game started, and there were some spectators standing on the northwest open platform. The match (Phillies versus Yankees) started at 7:30 PM and finished at 10:15 PM on the 7th of August, 2010. The artificial illumination after the match could fully satisfy the spectators' evacuation requirements. There were no difficulties in either way-finding or signage recognition. After the match, spectators immediately started to egress, and the vast majority of them chose to walk on the main ramp. The width and length of the ramp could facilitate many pedestrians and the continuous walking surface made walkers feel relaxed. Therefore, this main ramp was chosen as the observation and study site.

The evacuation process was video-recorded on site for 4 minutes, while a majority of spectators from the upper stands passed through the main ramp (Figure 2). We then analyzed video data in the laboratory to get individual flow data of each evacuee. Over 500 valid data plots on pedestrian movement and travel time were obtained. Crowd diffusion behavior on the ramp was also observed. When pedestrians walked towards the same direction, their desired speed was easier to achieve. The relaxed state encouraged greater differences in individual walking speed. In each time interval, pedestrians with higher walking speed showed overtake behavior. Also small group gathering in large crowds was noticed. Those family and/or friend groups tended to keep approximately the same walking speeds and closer personal distances, so that all group members could watch out for one another and stay all together along the way. These small groups of pedestrians often move at the speed of the slowest individual member in the group.

#### 3.2. Model Validation

Under the condition of \( \gamma_1 \) and \( \gamma_2 \) taking different values, the estimated results of flow rate in the downstream segment based on the diffusion model will be different as well. This study uses the minimum difference value between actual video sampling data on a number of pedestrians and estimated results in various time intervals through the downstream segment as the evaluation index to achieve the optimized \( \gamma_1 \) and \( \gamma_2 \) value as shown in

\[
\min f(\gamma_1, \gamma_2) = \sum_{j=1}^{t} \left[ q_A(j) - q_B^e(j) \right]^2,
\]
where \( J \) is the number of time intervals with pedestrian walk through downstream segment \( B \); \( q_B(j) \) is the actual number of pedestrians at time interval \( j \) in downstream segment \( B \); \( \tilde{q}_B(j) \) is the predicted number of pedestrians at time interval \( j \) in downstream segment \( B \).

Equation (8) shows that the prediction error indicator \( f \) is the function of \( \gamma_1 \) and \( \gamma_2 \). This study chooses \( \gamma_1 \) and \( \gamma_2 \) values which make \( f \) achieve its minimum value as the best value of these two parameters. The value intervals of these two parameters are \((0,1)\), starting from 0.1 and taking 0.1 as a time interval to determine the value of these two parameters. The set of value of each parameter is \((0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9)\). This set includes 9 values in total. By using these 2 parameter values to develop verification plan, there are 81 verification plans in total.

3.3. Result Analysis. As shown in Figure 2, upstream segment \( A \) and downstream segment \( B \) are chosen as study segments. The distance between these two segments is 100 m, \( L_{AB} = 100 \) m. The average pedestrian walking speed \( V \) is equal to 1.4 m/s. This is slightly larger than the average pedestrian walking speed on roads with a plane surface, because this passage was a downward walking ramp. The phenomenon of pedestrians surpassing people in front was commonly recognized during data collection and individual pedestrian walking speed varied greatly. Therefore, the level of diffusion in the pedestrian stream was relatively higher on the long passage ramp.

When predicting the numbers of pedestrians at each time interval in the downstream segment \( B \), the value of \( \Delta t \) is set as 5 seconds. The predicted results are shown in Figure 3.

As shown in Figure 3, the corresponding \( f \) value of scenario 34 is the smallest, only 0.03. This indicates that the corresponding \( \gamma_1 \) and \( \gamma_2 \) value of test scenario 34 could achieve the preferable prediction effect. Moreover, the prediction error is only 0.03 at this point. This indicates that the difference between the predicted number of pedestrians at each time interval through downstream segment \( B \) and the actual number of pedestrians was only 0.03 persons, which is a very small margin of error. Therefore, the proposed pedestrian diffusion model could accurately describe pedestrian macromovement behavior on the passage.

4. Conclusion

Diffusion behavior in pedestrian flow commonly exists in unidirectional crowd movement such as during the evacuation process of sports venues, transportation facilities, large public exhibition buildings, and pedestrians walking on roads and pavements in urban and neighborhood scale. Quantifying and describing such diffusion behavior more precisely could allow for analysis of pedestrian movement speed and density effectively in the above facilities, thus having an important impact on evaluating the circulation performance and egress safety of the facility planning and design. Meanwhile, the proposed model can be used in pedestrian simulation software as well for simulating unidirectional pedestrian macroscopic movement in buildings and urban environments and improving the accuracy and applicability of current pedestrian simulation tools.

In the modeling process, this study is based on the hypothesis that all pedestrian walking speeds obey geometrical distribution and discretize time to build the pedestrian flow rate prediction model for the downstream segments. The proposed model has the following advantages:
(1) the modeling structure is simple and could satisfy the requirements of real-time and rapidity; (2) the proposed model is a discrete model which is suitable for computer programming and is also easier to be embedded in pedestrian simulation software; and (3) obtaining a database does not require the movement trajectory of each pedestrian. The database only requires calculating the number of pedestrians in each time interval and the average pedestrian walking speed in the upstream segment, which could overcome the problem of complex structure in former macroscopic pedestrian simulation models.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (NSFC), Grants nos. 51308142 and 61304198, China Postdoctoral Science Foundation Funded Project (Grant no. 2014M550193), supported by “The Fundamental Research Funds for the Central Universities” (Grant no. HIT. NSRIF 2014072).

References
