ON THE SIMULATION OF PEDESTRIAN BEHAVIOR
A stochastic model based on Markov chain and information space

SHEN-GUAN SHIH, YEN-HUNG CHEN,
SHAN-CHING HU, CHING-YUAN LIN
National Taiwan University of Science and Technology

ABSTRACT: Pedestrians move according to their internal states and the spatial information that they perceive at their locations. We use Markov chain to define the state transition of a pedestrian as a stochastic process, and the concept of cellular automata to simulate the way information disseminates in space. We have implemented an agent-based computer program to observe some emergent behaviors from the interactions between pedestrians and the information space they are surrounded. It is expected that the representational model can be used to assist designers in analyzing the relation between building plan and the circulation pattern.

KEYWORDS: Pedestrian behavior, simulation

RÉSUMÉ : Les piétons se déplacent selon leurs états internes ainsi qu’en lien avec l’information spatiale qu’ils perçoivent autour d’eux. Nous utilisons la chaîne de Markov pour définir la transition d’un piéton comme un processus stochastique, ainsi que le concept d’automates cellulaires pour simuler la manière dont l’information se dissémine dans l’espace. Nous avons implémenté un programme informatique basé sur des agents pour observer certains comportements émergents de l’interaction entre les piétons et l’espace d’information qui les entoure. Nous anticipons que ce modèle de représentation puisse être utilisé pour assister les concepteurs dans l’analyse des relations entre le plan d’un bâtiment et le schéma de circulation.

MOTS-CLÉS : Comportement de piéton, simulation
1. INTRODUCTION

Pedestrian movement is an important research issue for the design, evaluation and management of building and urban spaces. Some researchers realized that pedestrian movement can be simulated with various models, including cellular automata (Yang et al. 2005; Schreckenberg et al. 2001; Fang et al. 2003), fluid dynamic (Helbing et al. 1992), multi-agents (Dijkstra et al. 2002; Haklay et al. 2001), axial maps (Kim et al. 2004), and magnetic forces (Matsushita et al. 1993). The focuses of research are mostly on the rebuild of pedestrian movement in a virtual world (Haklay et al. 2001) and/or the analysis of behavior patterns (Helbing et al. 2001).

Pedestrians move according to their preferences and the spatial information they perceive from their positions. A pedestrian finds the way to its destination in various means. One can look at signs, ask other people, use mental maps, or simply wonder around to see if any clue of the right way can be found. Such information may exist in various forms. We do not need to reconstruct the physical environment in details to simulate pedestrian movement, if we are able to specify the distribution of information in space. In this research, we assume a generic and simple pattern over the way spatial information disseminates through space by means of probability and statistics. Besides way-finding information, some physical factors of the environment such as temperature and noise may also influence pedestrian movement. Such factors can also be regarded as spatial information spreading out from every position to its surroundings. Various forms of information can be integrated into one platform so that a model based on the dissemination of information can be defined to setup the environment for the simulation of pedestrian movement.

A pedestrian may move towards a preplanned destination, or wonder for various attractions in a building or urban space. It would be difficult to define a representational model for human mind and derive sufficient data to initiate a realistic simulation. However, if we look from the macroscopic view at a large number of pedestrians, all kinds of specific influences to pedestrian behavior can be constrained into some distributional patterns of probability. This paper describes a model based on simulating the dissemination of information in space, and stochastic processes that define reactions of pedestrians who sense the information. The objective of the research is to devise a representational model, emphasizing the modeling of information dissemination in the space, and the internal state description of the virtual pedestrian. The research is based on the following assumptions. First, there are far too many factors for us to model in the environment that may influence the behavior of virtual human. However, it is possible to combine related factors, and define as some simple mathematical functions that describe the dissemination pattern of spatial information. Therefore the complexity of the real world can be largely simpli-
fied. Second, the variations of behavior of virtual humans might be large, but from the macroscopic view, the complex internal states of a virtual human can be defined with some stochastic models such as Markov process.

2. PROTOCOL ANALYSIS ON WAY-FINDING BEHAVIOR

To test the concept of pedestrians’ moving behavior in this research, we use protocol analysis to study way-finding behavior in an urban area. A complex underground shopping area which is intersected with passages to the subway, train station, and some city attractions such as department stores and hotels was chosen as the test place for the experiment. The way-finding behavior of seven subjects were recorded and analyzed. Each of the subjects was assigned to two way-finding tasks. All subjects were asked to think aloud, which is to talk to a voice recorder what they were thinking and trying to do as they tried to find their way to the assigned destination. A camera man followed each of the subjects to record the behavior along the way. The voice and video recording were analyzed and encoded for protocol analysis. Our analysis shows that the subjects iterate loops with similar structure. Each loop is consisting of four steps described as the following.

- Collecting information: seek for helpful information in the surroundings.
- Building mental map: integrate the gathered information with the mental map that is in mind. A mental map consists of information regarding spatial relationship among landmarks, path and destination, as well as the current position of self in that map.
- Decision making: Evaluate the current situation and decide what to do next. This step involves the evaluation of the mental map with the current situation to find if there is missing information or conflicts.
- Action: Take an action to alter the current situation for gathering more information or moving towards the destination. By action it means movements that lead to some change of the accessibility of information.

In most cases, from the voice and video recording we can observe clues for complete loops with all the four steps, and in other cases observable clues for one or two steps might be missing. The left-hand side of figure 1 shows a loop with the four steps as a circle, while the right-hand side draws the loop differently by breaking and stretching the circle along the time axis to show the behavior pattern.

The following table shows a snapshot from the recorded data and interpretation of one subject. The snapshot consists of four loops. Ambiguity and conflict are not totally avoidable in the interpretation but the overall looping structure seems to be apparent. Some missing steps might be the result of the subject’s incapability of thinking aloud all the time, and the incapability of video recording in showing some details of human behavior such as eye
movement. Therefore, we supposed that each step in the loop can be further broke down into smaller loops with similar structure. The finer loops could be observable with some other means of observation.

**TABLE 1. A PART OF THE DATA AND INTERPRETATION IN PROTOCOL ANALYSIS.**

<table>
<thead>
<tr>
<th>Time</th>
<th>Translated text from voice recording (in italic text) and observed behavior from video (in parenthesis)</th>
<th>Interpreted behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>05’54”</td>
<td><em>(look at the sign)</em> <em>I saw the sign, which is... So the Station underground street should go this way.</em></td>
<td>Collecting information</td>
</tr>
<tr>
<td>06’15”</td>
<td><em>(Read the map)</em> <em>(Let me take a look at the map.)</em> <em>(I am now at Exit 2.)</em></td>
<td>Collecting information Building mental map</td>
</tr>
<tr>
<td>06’33”</td>
<td><em>(Look around)</em> <em>(So I have to pass the train station hall all the way to the exit.)</em> <em>(Walk)</em></td>
<td>Decision making Action</td>
</tr>
<tr>
<td>06’45”</td>
<td><em>(Walk)</em> <em>(Taipei underground street turning right? Which way is correct? To the right? Or to the left?)</em> <em>(Look around)</em></td>
<td>Action Collecting information</td>
</tr>
<tr>
<td>06’59”</td>
<td><em>(Look around)</em> <em>(I see the Taipei underground street)</em></td>
<td>Collecting information</td>
</tr>
<tr>
<td>07’05”</td>
<td><em>(Look around)</em> <em>(There is a small passage, from which I can see the entrance of the underground street. Somehow it is still going down. I guess that the route splits underground.)</em></td>
<td>Building mental map</td>
</tr>
<tr>
<td>07’27”</td>
<td><em>(look around)</em> <em>(I am now at Exit 13 of Taipei underground street. Therefore...I suppose that Tien-Chien Restaurant should be...Wait, isn’t it the Taipei underground street? Or is it Chong-Shang underground street? Am I at the wrong street?)</em></td>
<td>Building mental map</td>
</tr>
<tr>
<td>08’11”</td>
<td><em>(look around)</em> <em>(So it is Chong-Shang underground street. I just saw the sign showing the direction to Chong-Shang underground street.)</em></td>
<td>Decision making</td>
</tr>
<tr>
<td>08’24”</td>
<td><em>(Turn around, walk back)</em></td>
<td>Action</td>
</tr>
</tbody>
</table>
3. THE MODEL AND SIMULATION OF PEDESTRIAN MOVEMENT

The experiment described above shows close relationship between spatial information, the mental map of the person, the decision making process, and the action he/she takes to alter the situation. A pedestrian may move towards a preplanned destination, or wander for various attractions in a building or urban space. In either case, the decision is made based on the internal state of the pedestrian and the information he/she perceives from surroundings. We concern possible reactions of a person when he/she perceives a certain pattern of spatial information. It is assumed that a pedestrian traverses within a finite number of states, each of which is defined by the target position he/she is heading for. Therefore, the simulation is divided into two tasks. The first is to define what kinds of spatial information are spreading out, and how. The second is to classify possible states of the pedestrian and the pattern of state traversing.

3.1. The dissemination of information

The simulation consists of a rectangular grid and a set of agents walking on the grid. A set of attractions are placed on the grid. Each attraction spreads out
the information of its existence to surrounding cells. Virtual pedestrians, collect information regarding the targeted attraction, and walk towards the direction which the information shows.

We use the term “pheromone” to name spatial information that may influence pedestrian movement, due to the analogy to ants moving around according to the distributional pattern of pheromone. Various kinds of pheromone spread out on the grid, depending on the content and purpose of simulation. Two kinds of information are considered in this research. The first type of information concerning way-finding in space. Each type of attraction for pedestrians in the space is defined as a source of pheromone. The pheromone disseminates from the source to surrounding areas according to some specific rules that define the amount of pheromone a cell receives from its neighbors. A pedestrian that is moving towards the attraction would follow the direction that is indicated by the strongest scent of the pheromone. The second type of information indicates physical condition of the space. Temperature, noise and crowd density are examples of the kinds of information they might be.

Pheromone disseminates into the surroundings in various ways, depending on the type of information and the type of grid cells. For example, we can use the shortest path algorithm to simulate the pheromone that represents way-finding information based on the assumption that the probability a pedestrian would find the right way to its destination is proportional to the distance of its current position and the destination. We do not need to simulate all the spatial characteristics regarding how the pedestrian finds the way if we are interested in only the macroscopic view of pedestrian movement. We can also use dissemination rules that comply with the physical law of heat transmission to simulate a map of temperature on the grid.

### 3.2. Pedestrian

A pedestrian is viewed as an information processing system that reacts to the accessible information. An agent that represents a pedestrian in the program is defined as a stochastic system with finite states. Each state of the agent represents a specific preference of the pedestrian to some attractions or places with desirable physical condition. Each state of the agent is associated to some rules that describe a preferred pattern of pheromone. The agent moves to reduce the difference between the preferred pattern of pheromone and that of the present position. When the current situation matches the desirable pattern, the action of the state is completed, and the agent switches to another state according to the specific possibility of the stochastic system.

Markov chain is used to define the probability of state transition of an agent. Using Markov chain, it is possible to combine various kinds of pedestrians into one single representation, despite of the fact that each pedestrian may have different preferences and purposes. It greatly simplifies the simulation for we
do not need to know the preferences of individual pedestrian. Instead we create a generic agent that represents all pedestrians in the simulation with one single matrix. The matrix that describes such a generic agent can be derived by a method called hidden Markov modeling (Giles et al. 1993). With hidden Markov modeling we can derive the unknown probabilities of state transition of a Markov chain from observable data. In this research we assume that for spaces of interest, we may derive one single state transition matrix for all pedestrians. For example, when we make a simulation on a hospital building, we may use only one type of agents, with one single state transition matrix to define the preferences of all pedestrians, including patients, nurses, doctors and administration people.

Preferences of pedestrians may change in simulation. Pedestrian movement would be different at the beginning of working hours, lunch time, and after working hours in a day. A simple way is using multiple Markov chains for different periods of time in a day. In this case, the discontinuity of pedestrian behavior upon the transition from one Markov chain to another can be smoothen by randomly selecting only a number of agents each time for applying the new Markov chain. The entire population would switch to another pattern of movement smoothly. Another solution is to define the state transition matrix as a function of time so that the transition probability varies with the time of a day and/or a week.

Agents may have interactions with other agents. Collision avoidance is a kind of short range interaction. Long range interaction may be simulated by defining agents as sources of pheromone, so that a crowd can be regarded as distraction or attraction in the simulation, depending on the state of the agent. Agents may have influences to the spreading of pheromone in the space and all agents would react to the dynamic pattern of pheromone.

4. THE SIMULATION

We have done some experiments using a prototype program implemented with the model described in this paper. At the time we are not able to compare the result of the simulation to realistic data of pedestrian movement. However, we do observe some interesting results from the behavior of agents in the virtual world.

We start from a very simple configuration. The space is divided into two areas with a wall in between them. On each side of the wall, two attractions are placed symmetrically. A number of agents are divided into two groups. One group moves alternatively between attractions 0 and 1, and the other group moves between attractions 2 and 3. The positions of attractions are marked with numbers in large font and the positions of agents are marked with small font in all figures showing the on-going status of simulation. The number that represents an agent shows the target attraction of the agent. After an agent has
successfully reached an attraction, it transits to another state stochastically with the state transition matrix of Markov chain. Agents in the experiments are initialized with various speeds. We assume a normal distribution of speeds with 40% of maximum difference in the simulated population.

**FIGURE 3. SIMULATION WITH 30 AGENTS FOR EACH GROUP WITH TWO ATTRACTIONS.**

Figure 3 shows a snapshot of a simulation with 60 agents divided into two groups of the same size. Some interesting result can be found in this simple experiment. First, the numbers of agents that are heading towards specific attractions are distributed equally regardless the initial distribution of the population. When the distribution of population is changed to some unbalanced proportion, such as all agents heading towards attraction 3 and zero for attraction 2, it restores to the equilibrium state within 200 steps in the simulation. As we can see in figure 1, the population of each group is restored back to 15 equally. The number of agents in a group is calculated with the average of 100 steps. Second, we find some agents do not choose the shortest path, which are the horizontal lines that connect two pairs of attractions. The agents that walk in the peripheral area are fast moving agents who detour from the shortest path to over pass agents that move slower, or simply avoiding collision with agents that head towards the opposite direction. Third, in some occasions, agents with the same direction line up to avoid frequent collision with agents heading towards the opposite direction, just like prior research on pedestrian movement has shown in their studies (Helbing *et al.* 2001).

**FIGURE 4. SIMULATION WITH 60 AGENTS IN EACH GROUP.**
In Figure 4, we increase the number of agents in each group to 60. This time the numbers of agents heading towards the same attractions are no longer distributed evenly. Although the spatial situations of all attractions are identical, the number of agents who head towards attractions 0, 1, 2 and 3 are 18, 41, 17, and 42 respectively. Notice that the proportions of population in the group of attraction 0, 1 and the group with attractions 2, 3 are roughly the same. And again, the state of equilibrium is stable regardless of the initial distribution of population. As we increase the size of population, the uneven distribution of population remain stable, but at a more extreme proportion, as shown in Figure 5. In these simulations, the attraction that has more agents heading towards may vary, but the proportion of population remains nearly constant.

**FIGURE 5. SIMULATION WITH 80 AGENTS IN EACH GROUP.**

Figure 6 shows that the proportion of population moving towards different attractions drops from 1 to less than 0.3 as group population increases from 30 to 100. Figure 7 shows the number of agents who reach the target attraction in 10 steps within groups of various populations. The figure shows that with 50 agents in the group, the circulation reaches its maximum efficiency. When the population is smaller than 50, the circulation does not reaches its full capacity, and when the population is greater than 50, circulation jam starts to hold back pedestrians from reaching the target attraction.

**FIGURE 6. PROPORTION OF POPULATIONS MOVING TOWARDS DIFFERENT ATTRACTIONS.**
With the simulation in figure 8, we tested another situation where positions of attractions are moved closer to each other. The numbers of agents heading towards attractions 0 and 1 are 18 and 12 respectively. The numbers of agents heading towards attractions 2 and 3 are 16 and 14 respectively. It is obvious that not only blocked circulation may cause uneven distributions; shorter circulation also causes uneven distributions.

**FIGURE 8. SIMULATION WITH VARIOUS DISTANCES BETWEEN ATTRACTIONS.**

Figure 9 shows a simulation of 60 agents in a building plan with a corridor of 1 meter in width. There is a circulation jam in the corridor. The total times of agents who have reached a target attraction in 10 steps is 66. After the corridor is widened to 2 meters, the circulation jam is lessened and the number of times an agent reaches a target attraction in 10 steps increases to 170, as it is shown in figure 10. The experiment shows that a designer can modify the design while the simulation is running, and see the effect of the modification very soon after the modification.
FIGURE 9. A BUILDING PLAN WITH A 1 METER CORRIDOR.

FIGURE 10. A BUILDING PLAN WITH A 2 METER CORRIDOR.

5. INTEGRATION WITH COMPUTER AIDED DESIGN APPLICATIONS

The prototype program can be integrated with computer aided design applications. The building plan can be modified without interrupting the simulation. In Figure 11, the entrance of the room in the center was moved to its upper right corner. Agents that are heading for attraction number 1 realize that change soon after the modification of the space layout and follow a new path to enter that room.

FIGURE 11. THE ENTRANCE TO THE CENTRAL ROOM IS MOVED TO ITS UPPER-RIGHT CORNER.
6. DERIVING DATA FROM REAL WORLD

A dentist clinic is chosen for collecting data for pedestrian movement. The floor area of the clinic is about 5 meters wide and 18 meters long. Four video cameras were installed in the clinic to monitor the space. The floor area of the clinic is divided into seven zones according to their usage. Figure 12 shows the plan dissected into zone A to zone G.

**FIGURE 12. ZONE A TO ZONE G IN THE FLOOR PLAN OF THE CLINIC.**

The numbers of pedestrians that visit each zone within every 10 second time span were counted. Table 2 shows a part of the observed data. Figure 13 shows the picture taken from the camera with boundaries of zones marked with dash lines.

**TABLE 2. NUMBER OF PEOPLE VISITING EACH ZONE WITHIN 10 SECOND SPANS.**

<table>
<thead>
<tr>
<th>Time</th>
<th>Zone A</th>
<th>Zone B</th>
<th>Zone C</th>
<th>Zone D</th>
<th>Zone E</th>
<th>Zone F</th>
<th>Zone G</th>
</tr>
</thead>
<tbody>
<tr>
<td>18:53:20</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>18:53:30</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>18:53:40</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>18:53:50</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td>18:54:00</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>18:54:10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>9</td>
</tr>
<tr>
<td>18:54:20</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>8</td>
</tr>
</tbody>
</table>

**FIGURE 13. A SNAP SHOT FROM THE MONITORING VIDEO.**
The collected data is analyzed using an algorithm called regression or curve fitting. State transition matrix that satisfies the observed data was estimated. However, the matrix is not unique, and the matrix with the minimum deviation from all data points could be the best guess. This desirable best-fitting matrix can be obtained by least square method, which uses the minimal sum of the deviations squared from a given set of data (Fraleigh et al. 1995). We used the derived state transition matrix for our prototype program to simulate pedestrian movement in the clinic. The picture in figure 14 shows a snapshot of the simulation, in which three people are staying in zone G, while one person is about to leave for zone B. There is a person near zone F is moving towards zone G. So far we have no measurement over the difference between the simulation and the actual situation.

**FIGURE 14. A SNAPSHOT FROM THE SIMULATION OF PEDESTRIAN MOVEMENT IN THE CLINIC.**

7. CONCLUSION

The result of the research is very positive to the objective of developing a tool for assisting designers to test pedestrian circulation in their design. It is expected that the model can be used to implement simulations that facilitate design process by providing early assessment to the circulation, as well as building management. The model can also be used to facilitate simulation games for education or for entertainment.

For realistic simulation, the model would require further calibration with data gathered from field study. With the current state of the program, it can be used to do virtual experiments for the discovery of agent behavior that may or may not have relations to pedestrian moving behavior in the real world. However, such virtual experiments can be used to do pilot study for designing realistic experiments or observation on real pedestrians.
ACKNOWLEDGEMENT

The research described in this paper is supported by the National Science Council of Republic of China, under the project NSC 97-2221-E-011-113-.

REFERENCES


