Modeling Planned and Unplanned Store Stops for the Scenario Based Simulation of Pedestrian Activity in City Centers

Jan Dijkstra and Joran Jessurun
Department of the Built Environment
Eindhoven University of Technology
Eindhoven, The Netherlands
e-mail: {j.dijkstra,a.j.jessurun}@tue.nl

***Abstract*—Micro-scale agent-based modeling can be used for the simulation of pedestrian movement for low and high density scenarios and of the effect of changes in an environment. Such models can also be used for pedestrian dynamics in city centers to show the design effects in the shopping environment. The main contribution of this paper is to introduce the implication of planned and unplanned store visits within a simulation framework for pedestrian movement simulation. This paper reports findings of planned and unplanned store visits by using Monte Carlo simulation.**

Keywords- Monte Carlo Simulation, Activity Agenda, Agent Based Modeling, Pedestrian Dynamics

I. INTRODUCTION

Agent-based modeling is a computational methodology that allows us to create, analyze, and experiment with artificial worlds populated by agents. A specific research area is micro-scale agent-based modeling that can be used for the simulation of pedestrian movement for low and high density scenarios and for the effect of changes in the environment. Such models can also be used for pedestrian dynamics in city centers to show the design effects in the shopping environment. In this context, Ali and Moulin [1] describe their multi-agent simulation prototype of customers' shopping behavior in a mall. Therefore, a multi-agent model to simulate pedestrian dynamic destination, route and scheduling behavior is under development, where the simulation of movement patterns is embedded in a more comprehensive model of activity travel behavior.

Representation is a main issue in simulating pedestrian dynamics. One can distinguish the representation of the pedestrian environment and the representation of pedestrians. In the domain of a city center, representation of a pedestrian environment includes the geometry of the shopping environment such as stores and streets, the network as a cellular grid, and pedestrian objects. Pedestrian representation includes socioeconomic characteristics, speed, goals, familiarity with the environment, and activity agenda. It is assumed that pedestrians perceive their environment and that they are supposed to carry out a set of activities. For completing an activity, pedestrians spend time in stores. As a consequence, time duration influences their movement behavior over the network.

Although a 3D presentation of pedestrians and the pedestrian environment for the simulation of pedestrian movement is the ultimate goal, it is nevertheless meaningful to test the underlying principles in an appropriate 2D representation of pedestrians and their environment. NetLogo can be used as a simulation toolkit because it is a suitable simulation framework that supports modeling, simulation and experimentation. It also offers skeletons of agents and their environment, and interoperability (e.g., Geographic Information System (GIS)). We will use shape-file information of the environment and network structure for visualizing the 2D environment and NetLogo for the actual simulation.

The main subject of this paper is to introduce the implication of planned and unplanned store visits within a simulation framework for pedestrian movement simulation. This framework involves an agent-based model that provides an activity agenda for pedestrian agents that guides their shopping behavior in terms of destination and time spent in shopping areas. In order to implement the activity agenda, pedestrian agents need to successively visit a set of stores and move over the network. It is assumed that pedestrian agents' behavior is driven by a series of decision heuristics. Agents need to decide which stores to choose, in what order and which route to take, subject to time and institutional constraints. It is assumed that pedestrian agents are in different motivational states. They may at every point during the trip have general interests in conducting particular activities, without having decided on the specific store to visit, but they may also be in a more goal-directed motivational state in which case they have already decided which store to visit. The motivational states are of influence on the impulse and non-impulse store choice processes and therefore on the planned and unplanned visits to a store. All these aspects affect pedestrian agents' time duration in visiting stores. Pedestrian agents move over a street network and are part of a pedestrian flow in this street network. However, pedestrian agents can be temporarily removed from the pedestrian flow by visiting a store and participating again in the pedestrian flow after visiting that store. In that case, the time spent by a pedestrian agent in a store is relevant. For the simulation run this time duration is determined by a Monte Carlo simulation [2]. In this paper, the focus is on the number of planned and unplanned store visits because that determines the activity agenda. Successful completing an activity by a store visit influences the remaining planned and unplanned store visits. The findings from the collected data of the number of planned and unplanned store visits indicate that this number meets the Gamma distribution, and that this number also depends of

ISBN: 978-1-61208-308-7

the motivation and some socio-economic characteristics of the visitor to the city center.

This paper discusses successively pedestrian movement simulation in section II, simulation process in section III, and planned and unplanned store stops in section IV. A discussion about the conclusions and future directions will conclude this paper in section V.

II. PEDESTRIAN MOVEMENT SIMULATION

Pedestrian movement simulation consists of pedestrian agents and a simulation environment, which consists of a street network and a set of stores. Polygons are used to indicate borders and functional areas such as walkways. Each cell in the network has information about which agents and polygons occupy it. Also, it contains information about other features such as appearance of stores that are observable from that cell. A pedestrian agent moves with his own behavior and personal characteristics. Every time step, there is an update about agent's positions. The cellular network provides percepts to the pedestrian agent and the pedestrian agent performs actions based on their percepts. Behavioral principles drive the pedestrian movement. Details of the equations representing these behavioral principles are beyond the scope of this paper and are presented in Dijkstra et al. [3], but are related to the *Agent Loop* in Fig. 1. They include the perceiving of the environment, the possible match of the percepts with the activity agenda and as a consequence the determination of the activation to a store with as a result a completed activity with a consequence for the activity agenda.

In the case of the test ground of the city center, each store consists of a cell containing store information. A street network consists of cells with cell information. Pedestrian agents are situated in the cells of the network, namely a street cell or a store cell. The network is irregular because a clear border between a store and adjacent cell is desired. Additionally, each cell in this network is identified by its node and these nodes are linked together.

For populating pedestrian agents in the environment and for attaching activity agendas to pedestrian agents, a Monte Carlo simulation is used which implies that the behavior of each pedestrian agent is simulated by a series of draws of random numbers from successive probability distributions [4]. These probability distributions are based on real data collections, such as time spent in a store, attaching inner lane or outer lane as an entry point, speed, and pedestrian characteristics (gender, age, etc.).

NetLogo is used for the simulation because it easily allows the empirical testing of the principles of the simulation approach. An attractive feature is its ability to integrate GIS data directly into the simulation. With the integration of GIS, pedestrian agents can move around 'real space'. In the simulation, MapInfo data will be integrated with NetLogo. On the basis of this GIS, a network structure with nodes and links will be generated. The nodes are related to polygons in the original drawing, and the links will be determined by the topology of the polygons. The information of the polygons is available for pedestrian agents moving in the network. This information could be store-related information, but also information about the area and perimeter of the polygon. Fig. 1 shows the activity diagram of the simulation setup. The simulation process starts with loading the environment involving GIS information and databases for instance activity agendas and personal characteristics for creating pedestrian agents. The creation of an initial situation at time t_b (beginning time) means that the environment will be populated with pedestrian agents. The simulation run starts at time t_b. The simulation time step includes the creation of zero or more pedestrian agents using the Monte Carlo method: a pedestrian agent would need to be assigned an initial scenario. Also, there is an update of pedestrian agent scenario's that results in pedestrian agent actions and a schedule of the next step.

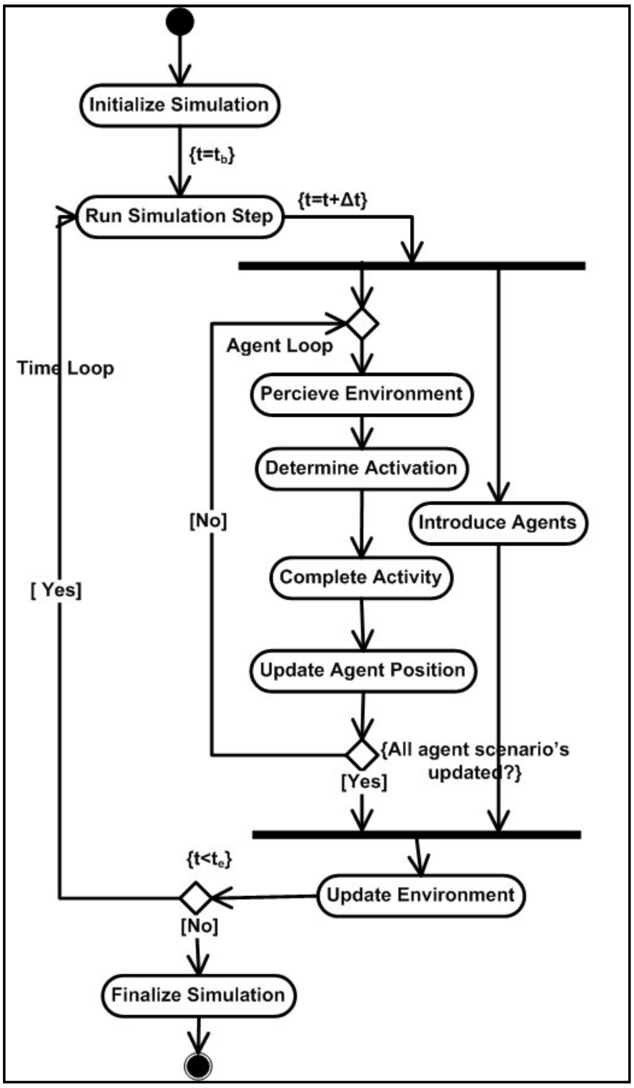

Figure 1. Activity diagram of the simulation setup.

The consequence is the movement to a new position. Then, an update of the environment will be realized. The simulation run stops at time t_e (ending time).

III. SIMULATION PROCESS

This section provides some understanding in the engineering basis of the simulation process using NetLogo [5]. The model structure is based on contexts and projections [6]. The core data structure is called a context that represents from a modeling perspective an abstract population; the objects in these populations are referred as agents. The context provides the basic infrastructure to define a population and the interactions of that population; it creates an abstract environment in which agents exist at a given point in the simulation. The context also holds its own internal state for maintaining the collection of agents. This state can consists of multiple types of data. These provide agents with information about the world in which they interact. In addition, data fields can be maintained by the context. A data field is an n-dimensional field of values with which the agents in a context can interact. These data fields can be directly associated with a physical space. The field is generic, which means each value is derived from a set of coordinates.

Projections take the population as defined in a context and impose a new structure on it. This structure defines and imposes relationships on the population by using semantics defined in the projection; therefore an agent population is realized once a projection is applied to it. This means that projections are added to a context to allow agents to interact with each other. Each context can have an arbitrary number of projections associated with it (1-n relationship); in our case it concerns about two projections.

A feature of NetLogo is the ability to integrate GIS data directly to the simulation; it provides a set of classes that allow shape-files to be displayed. For example, shape-files can be provided by GIS software packages like MapInfo and ArcGIS. A GIS contains multiple layers of data; each layer is made up of a number of elements. Each feature in the layer has two aspects to it, its geographical coordinates (but it could be also a polygon, polyline or polypoint) and the data associated with it. GIS store data about layers in database files, with each record in the file referring to a feature in GIS. Actually, NetLogo integration with GIS means shape-file integration while they use the same shape-file; NetLogo is used to read the shape-file data. Agents are created using these data and the simulation process. This means that the context creator provides the population. Agents can be created, re-created and destroyed at every simulation step. The interaction with the environment is provided by the shape-file containing GIS data; multiple GIS layers are the projections within NetLogo. Two projections are assumed, one for the GIS data and the other one for the generated network from this GIS data. The context needs these GIS data for the data fields which provides the information from the environment.

In this approach, the environment consists of polygons representing the network of stores and streets. In fact, this network is divided into cells, namely *store* cells and *street* cells. Each cell is identified by its node. For instance, pedestrian agents can move from a *street* cell to an adjacent *store* cell. Cells containing store information are not always strictly adjacent, for example. In a GIS software package, feature data will be connected to cells of the network and layers will be created. After that, the GIS software package provides the shape-file that will be loaded in NetLogo. This shape-file provides the environmental information. The simulation run can be performed. Each cell in the network has a node. An agent is located in a node on the underlying representation and can move on the implicit generated network to other nodes. Strictly speaking, it does not follow a cellular automata approach because an agent moves from node to node and is situated randomly in the cell related to that node.

The test ground is the inner-city center of Eindhoven. The simulation will be performed on a part of this city center, particularly on a section of the city center. Fig. 2 shows the cellular grid of this segment; Fig. 3 shows a possible population of pedestrian agents in a part of this segment, and Fig. 4 shows the nodes in this part of the segment.

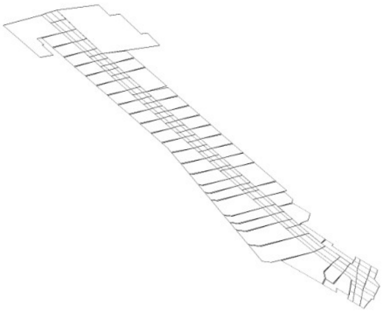

Figure 2. Segment of as section of the inner-city center of Eindhoven.

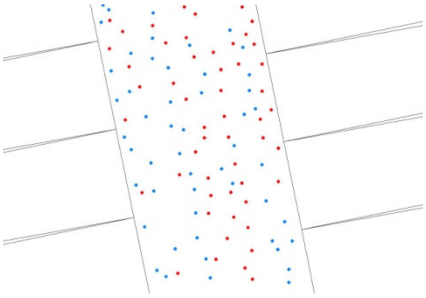

Figure 3. Population with agents in a segement part.

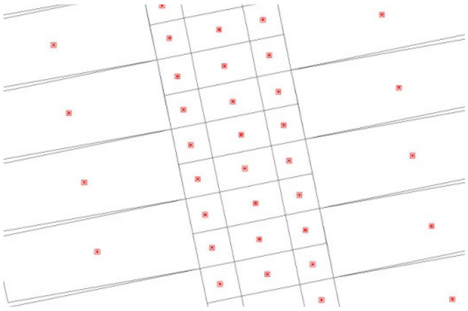

Figure 4. Nodes in this segment part.

TABLE I. POLYGON FEATURES OF THE NETWORK

Identification	**Type**	**Category**	**Description**	**Priority**
12002	0	5	McDonalds	1.00
12004	0	3	Etos	1.05
11001	0	1	C&A	2.30
13001	1	0	C&A	0.00

Table I shows a number of features as part of the provided MapInfo GIS database; each cell in the network includes these features such as an identification number, type indicating store cell or street cell, and category indicates store category (for example category number 3 represents a health&body store). This database is used by the data fields that are used in NetLogo for environmental information for pedestrian agents perceiving this environment. Priority means the proportion of visits; for example a priority of 2.3 means that 2.3% of the visits are intended to visit that store.

The pedestrian agent model system simulates which shopping activities are conducted by pedestrians, where (destination choice), when (choice of timing), for how long (duration), and which route is used for implementing the activity agenda (route choice). All these activities influence pedestrian's positions in the simulation run. Data collecting efforts are needed to calibrate the agent model system for the test ground: a survey that is a sample of respondents who are asked about their activity agenda. This survey includes questions for pedestrians who have completed their visit to the city center and ask them about the nature of their completed activity patterns (which store, for how long, sequence, and route). Also, a survey was conducted about pedestrian's awareness of stores and signaling intensity of stores as well as the visit of a store and the completion of activities.

IV. PLANNED AND UNPLANNED STORE STOPS

Borgers and Timmermans [7] used Monte Carlo simulation for the incorporation of the numbers of stops and the sequence of planned stops/purposes, because in their opinion the concept of multi-stop, multi-purpose behavior is relevant for understanding pedestrian behavior. According to this line of thought, we assume an activity agenda includes a number of planned and unplanned store visits that can also be considered as a number of non-impulse and impulse store visits.

Every pedestrian agent receives at its introduction in the simulation a pedestrian scenario. This pedestrian scenario includes besides general characteristics like gender, age, companionship also familiarity with the city center, motivation, time budget, and activity agenda. After a store visit the activity agenda will be rescheduled. The number of planned and unplanned store visits is determined by a Monte Carlo simulation.

For this purpose, data from visitors to the city center of Eindhoven are gathered by interviewing them about their motivation and the stores they visited. They are also asked about successful visits and which of them were planned and unplanned; 402 routes are identified.

The findings from the collected data of the number of planned and unplanned visit stops show a skewed distribution. The skewed distributions are different depending of gender, age category and motivation. We often need a skewed distribution where probability densities below and above the mean are distributed differently. In this case, we assume, by analogy with multiple stops, a Gamma distribution.

The probability density function of the Gamma distribution is given by:

$$f(x;k,\theta) = x^{k-1}\frac{e^{-x/\theta}}{\theta^{k}\Gamma(k)} \qquad (1)$$

where, k is the shape parameter (k>0) and θ is the scale parameter (θ>0). Both k and θ will be positive; they are derived from the skewed normal distribution of the number of (planned) stops from their data collection depending on gender, age category and motivation. The shape parameter k is derived from the skewness of the skewed normal distribution (see Fig. 5 for an example) and the scale parameter θ is derived from the mean and k; these parameters are given by:

$$k = \left(\frac{2}{skewness}\right)^{2} \;,\; \theta = \frac{mean}{k} \qquad (2)$$

Table II shows the values of the parameters of the Gamma probability distribution for different motivation, gender and age category. From the collected data, for the age category was only possible to distinguish between over 55 and less than 55 years.

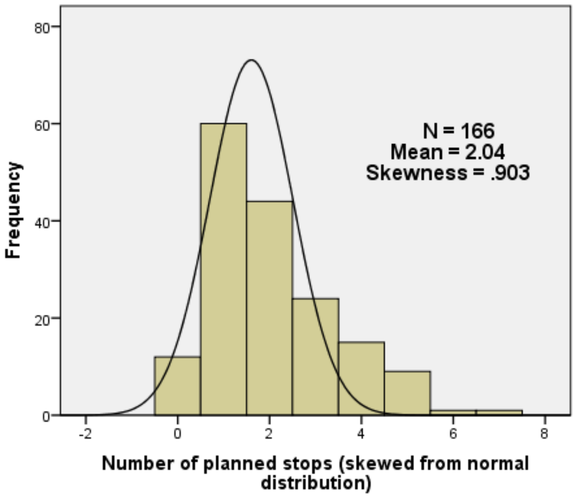

Figure 5. Number of planned stops; Goal Oriented motivation, female & age<55.

TABLE II. PARAMETER VALUES GIVEN DIFFERENT CATEGORIES

Category	Parameter	Motivation				
		Goal oriented		*Leisure oriented*		*No spec. intention*
		Num. stops	Plan. Stops	Num. stops	Plan. Stops	Num. stops
Man						
< 55	*k*	2.079	2.419	11.973	2.356	4.655
	θ	.957	.662	.279	.845	.421
≥55	*k*	2.773	1.821	64.515	28.597	1.333
	θ	.721	.890	.035	.050	1.252
Female						
<55	*k*	4.331	4.906	711.11	9.467	110.80
	θ	.623	.416	.006	.261	.028
≥55	*k*	1.897	1.174	31.210	95.181	5.642
	θ	1.270	1.678	.138	.025	.507

Striking is the distinction between the number of stops and the planned number of stops. The more goal oriented the less there are unplanned stops.

The Gamma inverse function *G(p)*, which is the inverse cumulative distribution function, is given by (3). Given a random number *p* from a uniform distribution in the interval (0, 1), the value of *G(p)* has a Gamma distribution with parameters *k* and *θ*. That means, given a number p on the x-axis provides the number of stops on the y-axis; where real values are rounded to integer values.

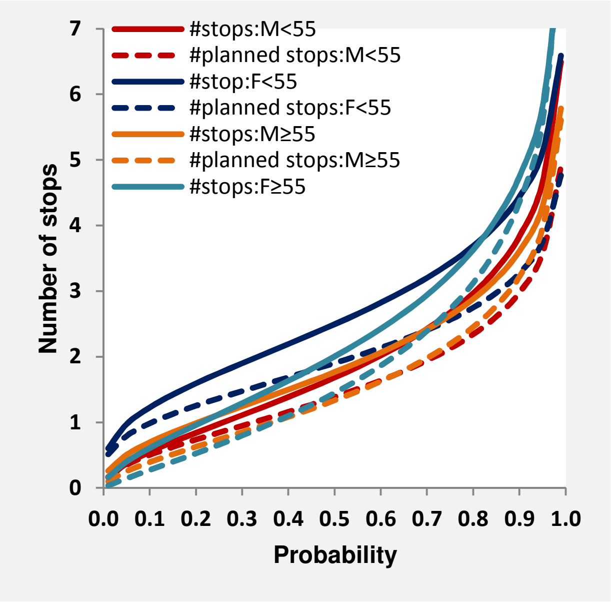

Figure 6. *G(p)* for Goal Oriented motivation, different gender and age category; Number of (planned) Stops vs. Probability.

$$G(p) = F^{-1}(p;k,\theta) = \{x: F(x;k,\theta) = p\}$$

where (3)

$$p = F(x;,\theta) = \frac{1}{\theta^k \Gamma(k)} \int_0^x t^{k-1} e^{-x/\theta}$$

Fig. 6-8 show the *G(p)* distribution for respectively *goal oriented* orientation, *leisure oriented* orientation and *no specific intention* orientation with respect to the number of (planned) stops.

Also there is a distinction for gender (male, female) and age (<55, ≥55 years). The number of unplanned stops can be derived from the calculation of the number of stops minus the number of planned stops.

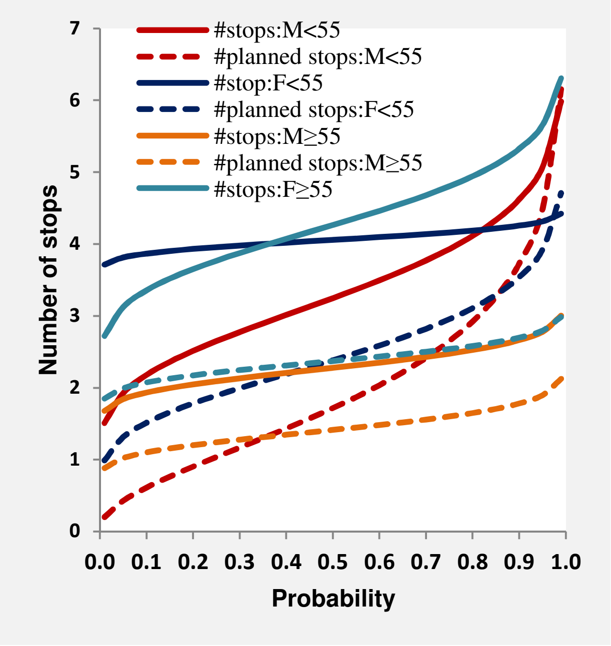

Figure 7. *G(p)* for Leisure Oriented motivation, different gender and age category; Number of (planned) Stops vs. Probability.

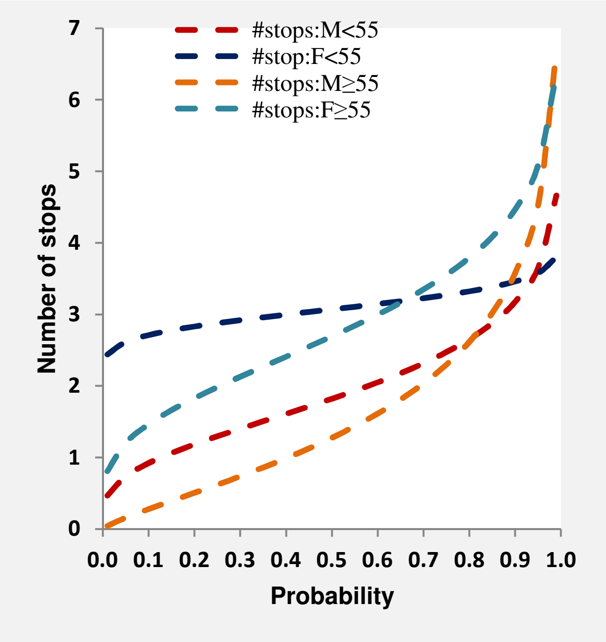

Figure 8. *G(p)* for No Specific Intention motivation, different gender and age category: Number of Stops vs. Probability.

What is clear form Fig. 6-8, that goal oriented motivation more corresponds to a planned route. Leisure oriented motivation more corresponds to an unplanned route, and if one has no specific intention then one has an unplanned route and no planned stops.

V. DISCUSSION AND FUTURE DIRECTIONS

In this paper, we presented only a small part of a simulation platform context for performing the pedestrian movement simulation, namely planned and unplanned visits. Pedestrian behavioral principles were briefly mentioned, and the pedestrian movement simulation setup as well as simulation process was briefly discussed, shown in the activity diagram of Fig. 1. In the current state of the pedestrian agent model system, which is the simulation, all data for the pedestrian movement simulation were collected. The findings about the number of planned and unplanned visits to a store were presented. This gives the number of planned and unplanned stops for the activity agenda in the simulation process. The framework for processing agent-based pedestrian activity simulations will be implemented and all the data will be integrated in the agent model system. This is a step by step process. At this moment, the signaling intensity of a store is represented by a priority of visiting the selected regarding store; and the store visits are split up in planned and unplanned visits.

The pedestrian model system will be tested in a 2D environment, because we want to validate the basic principles. Also, pedestrian agents move from node to node and are situated into the cells related to those nodes. They are situated randomly in those cells, but if the cell is completely occupied by other pedestrian agents, they cannot move to that cell. This approach reduces the complexity of the simulation by ignoring collisions and with that collision detection. These certain characteristics of the system make the simulation feasible because computer power is less binding. If all the parts are implemented, validation of the pedestrian model system will be performed. The data collection will be split up into two parts and the results of the separate simulation experiments will be compared.

Future developments should make the pedestrian agent model suitable for a 3D environment with lifelike virtual persons. In that case, the pedestrian agent movement will be realized from cell point to cell point considering collision detection. Finally, this will result in a virtual environment of a real situation, populated with virtual persons and a real person (user) moving amongst these virtual persons. A user can assess an environment that has high reality content. Preferences of users can be collected and the utility of a proposed situation can be estimated where appropriate. With this approach, it is possible to gain a deeper insight into the activity behavior of city center visitors and thus in the pedestrian flows in city center environments, even for those that do not exist yet.

REFERENCES

[1] W. Ali and B. Moulin, “How artificial intelligence agents do shopping in a virtual mall: a ‘believable’and ‘usable’ multi-agent based simulation of customers’ shopping behavior in a mall”, in Canadian AI, LNAI 4013, L. Lamontagne and M. Marchand, Eds. Berlin: Springer-Verlag, 2006, pp. 73-85.

[2] J. Dijkstra, A.J. Jessurun, H.J.P. Timmermans, and B. de Vries, “A framework for processing agent-based pedestrian activity simulations in shopping environments”, Cybernetics and Systems, Vol. 42, No. 7, 2011, pp. 526-545.

[3] J. Dijkstra, A.J. Jessurun, and H.J.P. Timmermans, “Simulating pedestrian activity scheduling behavior and movement patterns using a multi-agent cellular automata model”, in Proceedings of the Transportation Research Board Conference, Washington, January, 2002.

[4] M. Bierlaire, G. Antonioni, and M. Weber, “Behavioral dynamics for pedestrians”, in Moving through Nets: the Physical and Social Dimensions of Travel, K.W. Axhausen Ed. Elsevier Science Ltd., 2005, pp. 81-105.

[5] W. Zhu and H.J.P. Timmermans, “Cut-off models for ‘go-home’ decision pedestrians in shopping streets”, Environment and Planning B: Planning and Design, Vol. 35, No. 2, 2008, pp. 248-260.

[6] R. Najlis and M.J. North, “Repast for GIS”, in Proceedings of the Agent 2004 Conference on Social Dynamics: Interaction, Reflexivity and Emergence, C.M. Macal, D. Sallach and M.J. North Eds. Chicago, Illinois, 2004, pp. 225-260.

[7] A. Borgers and H.J.P. Timmermans, “City center entry points, store location patterns and pedestrian route chocie behavior: A micro-level simulation model”, Socio-Economic Planning Sciences, Vol. 20, 1986, pp. 25-30.