Pedestrian Behaviour Modelling

An Application to Retail Movements using a Genetic Algorithm

KAY KITAZAWA and MICHAL BATTY

Centre for Advanced Spatial Analysis, University College London

Key words: Pedestrian modelling, Retail movement, shortest-path, Genetic Algorithms

Abstract: This paper introduces a study of pedestrian behaviour modelling which incorporates ideas about agent-based systems and the traffic models based on the utility-maximization theory. The aim of this study is to examine the applicable scope of one of the basic assumptions in existing spatial behaviour models; the route with shortest distance maximizes the utility of each pedestrian's travel. Although shortest-path models have been widely used in the field of Traffic management to predict routing behaviour, there can be seen a lot of erratic behaviour in urban areas, shopping migration behaviour for instance, which can not be explained by them. Thus, it is important to identify other possible influential factors on their utility maximization process in order to develop more explicable models of pedestrian movements. In this study, we implemented a simulation model using the shortest-path model as one of evaluation criteria of Genetic Algorithms (GA) to computationally emulate retail movements of shoppers in a big shopping centre and to test the accuracy of the model by comparison between the routes estimated by the model and actual trajectories of shoppers. This simulation system will be used as a platform for further modelling.

1. INTRODUCTION

Pedestrian movement in town centres has attracted not only urban planners or government officials but also retailers, advertising agents and those who are involved in management of urban space. Although their prime objectives or needs are varied, their particular interests are in measuring and modelling pedestrian movement for understanding the way which people move in an urban setting as well as assessing the vitality and viability of town centres. This paper outlines an agent-based model which is intended as the beginning of such predictive models and describes a simulation using Genetic Algorithms (GA) to recreate retail movements of shoppers in a large shopping centre.

We start off with the social background of the needs for modelling pedestrian movement in town centres in section 2 and then outline some previous approaches to understanding and predicting spatial behaviour. Our brief survey suggests that only a small number of researches have been carried out in this field and that there has been no methodology established yet which can describe pedestrian's erratic movement in micro-scale environment. In section 3, we move on to a tentative framework of a new pedestrian behaviour model, which incorporate agent-based modelling techniques and utility-maximization theory. Section 4 describes a simulation model as the first step of implementing this framework. In this simulation, shopper's trajectories are estimated by GA system with shortest-path as its evaluation function. The aim of this simulation is to examine the applicability of one of the basic assumptions in existing models of spatial behaviour; the route with the shortest distance maximizes the utility of each pedestrian's travel. The remainder of the paper discusses the results of the simulation and potential factors which influence pedestrians' movement.

2. BACKGROUND

The concept "sustainable development" and "compact city" has increased social-economic importance of town centres, especially those for pedestrian. A number of local authorities and communities have started taking the initiative in implementing pedestrian-oriented policies not only in urban planning but also in traffic management and other economic strategies in order to create and maintain more liveable town centres. Greater London

Authority (GLA), for example, set it as one of their central planks to improve current town central pedestrian's environment into more walking friendly one by promoting use of more sustainable transport modes (Livingstone, 2001). The GLA also works in partnership with boroughs and other relevant agencies to develop a town centre network in London, which consists of dispersed accessible nodes of services and consumer choice across London (GLA, 2004). In recognition of the importance of high streets and town centres as a shopping centre as well as a place to live and work, several local councils in London such as The Royal Borough of Kensington and Chelsea and London Borough of Ealing have carried out major programmes of improvements of such public spaces. These programmes include renewal of pavements, improvement of public transport links, road crossings and signage systems as well as provision of good quality street furniture and facilities (The Royal Borough of Kensington and Chelsea, 1997). The aim of these programmes was to attract more people to live, shop, work, learn, relax and invest in these places in order to maintain and enhance the vitality, viability and attractiveness of the town centres (Centre Partnership and London Borough of Ealing, Regeneration & Major Project Unit, 2002).

2.1 Needs for modelling pedestrian

Despite recent pedestrian-oriented trends in urban planning, the methods used in on-the-ground planning processes are often inadequate and insufficient to predict patterns of pedestrians' spatial movement. "Research or detailed investigation on behaviour of pedestrians has been neglected in transport management" (Gemzøe, 2001). This is because that most transport modelling until recently exclusively focused on motorised transport as urban transport system has been highly dependent on vehicular traffic. Urban planners who want to implement walking-friendly urban environment schemes need new tools to help them with understanding present situations of pedestrians in town centres by measuring flows, analyzing their patterns, and identifying the functioning of movement systems. One of the key requirements for these tools is to provide accurate estimations of the impact of proposed plans on pedestrian behaviour. There have been growing needs for new behavior models of pedestrian which can predict how people move around in urban central areas.

2.2 Current spatial behaviour models

In order to meet these needs, a statistical model which predicts total walking volumes (average number of people per hour) for every node on the street network in Central London is now being developed by Transport for London (TfL), the Central London Partnership (CLP), and Intelligent Space Partnership (ISP). The approach adopted in the model is the standard Multiple Regression Analysis and it infers the relationship between a series of independent variables such as visibility (Maximum Radial line of sight) within the street network, accessibility to an underground station, pavement width, land-use, on a dependent variable, pedestrian flows (Desyllas, Duxbury et al., 2003). General structure of the model fits the observed data with a strong r-squared value and it also suggests that visibility is the most influential factor on walking volumes.

There have been several aggregate models similar to this for describing people's spatial behavior such as Huff model and Marcov-chain model. Although they have been widely used mainly in the marketing research field and proved to be effective enough to estimate a potential attraction of a new shop, or to predict the approximate number of visitors or probability that a certain place will be chosen as someone's destination, they can not deal with pedestrian movement at the scale of buildings and streets (Kitazawa et al. 2003). As these models are designed to be applicable at an urban-wide scale, they have rarely been successfully applied to modelling microscopic behaviour such as flocking and crowding. For the purpose of modelling such behaviour, not general patterns of movement, models which are capable of recreating movement of individuals are strongly needed.

Multi-agent-based micro simulation approaches in line with the idea of complex system arose as alternative modelling techniques. This type of model is one in which the basic unit of activity is the agent. Each agent autonomously acts according to its behaviour rules to achieve its goal. The outcomes of the model are determined by the interactions of the agents. TRANSIMS model combines traditional transport gravitation models and socio-economic data of zones that create traffic demand and a real-time agent-based simulation of the activities of individual travellers (Nagel Beckman, et al., 1997). STREETS model took the same approach but focused more on activities of pedestrians in urban districts (Schelhorn, O'Sullivan et al., 1999). The behaviour rules that pedestrian agents follow are determined by spatial configuration, pre-determined activity schedules, the distribution of land-uses, 'vision' of agent which enables the agent to search and recognise buildings near its route. Although this technique has been adopted in a number of areas, it may be the research of evacuation in which applications of this technique have most widely been used.

Myliad/Simulax model (Thompson, 1994) determines critical parameters such as level of service by unit area, flow rates for normal and emergency egress and simulates total evacuation times as well as individual's evacuation route. However, most of these multi-agent-based simulation models deal with only quite simple events in which people act under few behaviour rules such as avoiding obstacles, to attain the single objective. Real-life microscopic pedestrian movements, however, are much more complicated and often accompany seemingly erratic behavior. What brings this complexity and unpredictability to pedestrian behavior is perhaps that each pedestrian is apt to change his/her original destination or even purpose(s) of the behaviour itself (Ishibashi, Kumata et al., 1998). Thus it is important to take into account the functioning of people's selection behaviour, in other words, the way in which people make a choice.

Several pedestrian modelling approaches which describe discretionary activities such as route choice behaviour have been proposed and tested. Discrete-choice-based models which work in accordance with the potential function model (Hughes, 2002) have been proposed (Hoogendoorn, 2003, Hoogendoorn and Bovy, 2003). In these models, pedestrian route choice behaviour is described as a series of a subjective rational choice between infinite alternatives and "pedestrians are assumed to optimize their decisions given constraints from his or her activity agenda and risks involved in their decisions, while taking into account the uncertainty in the expected traffic conditions… This uncertainty reflects to among other things lack of experience, observability and randomness of future conditions, and thus pertains to non-deterministic route characteristics" (Hoogendoorn, 2003). The framework of the model which is based on utility-maximization proved to be applicable and effective to model human decision-making.

2.3 The scope of this study

Therefore this study aims at developing a generic model of pedestrian spatial behaviour which incorporates ideas about agent-based systems and the discrete-choice models based on utility-maximization theory. Since it is beyond the scope of this paper to implement the entire model framework as a simulation system, we, at this stage, develop a simple simulation using Genetic Algorithms. The purpose of this simulation is to examine the applicable scope of one of the basic factors in existing route choice behaviour models; the shortest path. Through comparison between the results of the simulation and observed data from the surveys of retail behaviour in a shopping centre, we try to identify other influential factors on pedestrian movement for future modelling.

3. NEW MODEL FRAMEWORK

A tentative framework of pedestrian behaviour model on the basis of multi-agent concept in conjunction with discrete-choice models based on utility-maximization theory has already been suggested (Kitazawa, Shibasaki, et al., 2003). It includes transition of the aims of the journey and feedback-loop of information-gathering process which leads to the transition. In this study, we develop a generic model in line with this framework. The basic structure and concept of this model is shown in *Figure 1*.

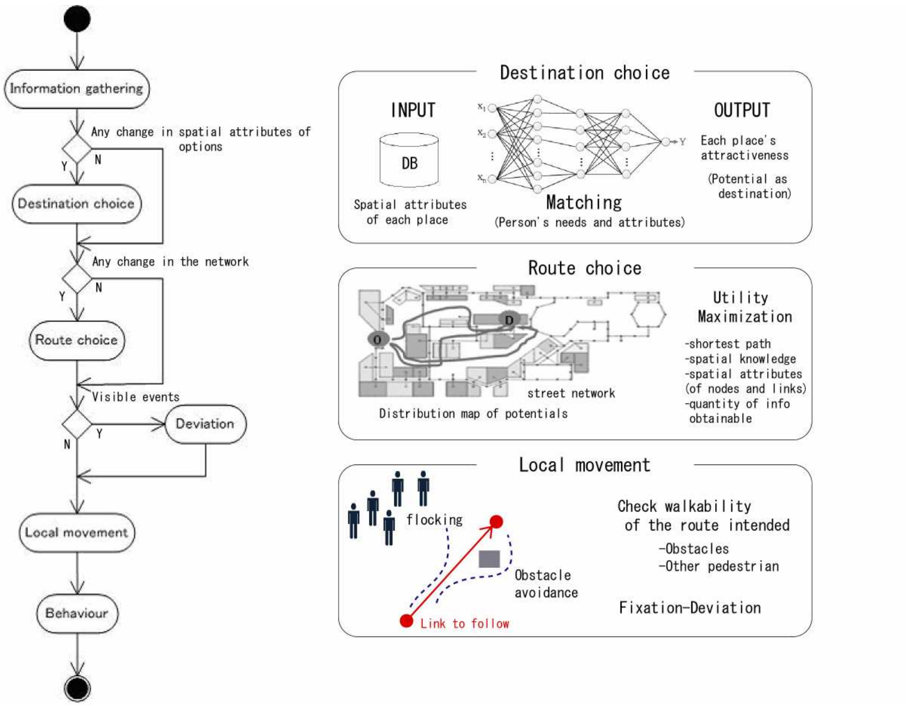

Figure 1. Concept of a new pedestrian behaviour model

Initial loading of the simulation with pedestrian agents with different attributes and pre-set activity schedules is achieved using socio-economic data sets. The attributes of pedestrian agents include their "physical" strength such as walking speed and fatiguability as well as general tastes/preferences which are related to their income, gender and age. These preferences are used to set a tentative activity schedule that is a sequence of places (nodes in the street network) which the agent intends to visit during their travel. Given these characteristics and schedule plan, each agent decides their next behaviour (movement) in four disparate processes.

1. Information gathering
2. Destination choice
3. Route choice
4. Local movement

The first process is information gathering. Each pedestrian gathers information about the attributes of places which are possible destinations, the street network and other pedestrian and then compares it with that from the previous states. The system proceeds to the appropriate processes, either process 2 or 3, if there is any change found, otherwise to process 4. For example, if we simulate retail behaviour assuming that shop *A* opens from 9 to 17, one of the attribute of shop *A* "availability" will change at 17 o'clock. As it discourages the shoppers who intended to visit the shop from going there, they need to revise their activity schedule when they get the information. Similarly, pedestrians need to find alternative routes when they noticed that one of the links of their intended route is closed or congested.

In the destination choice process, the system matches general tastes/preferences of pedestrians with the attributes of places such as size of the premises, type of customer that the place targets. This process, using marketing data in conjunction with neural network algorithms, identifies the attractiveness of each shop to each pedestrian and its results will be provided in the form of a probability of each place being chosen as a destination. Other attributes such as availability and travelling cost (distance from present location) are also taken into consideration. A potential distribution map can be drawn by these values and then the places of which potential is above the threshold will be set as intended destinations.

Next, the route choice process selects the "optimal" route which connects all intended place to visit under the time constraints from the street network data. This process will be refined by introducing travelling salesman algorithms with several optimization criteria such as shortest path, width of the streets (links) and the number of attractive places to which each street faces. It will also be enhanced by changing parameter values of optimization according to the different levels of spatial knowledge of each pedestrian. We will use Mixed Logit model as the basis of this optimization. This enables us to introduce a probability distribution that reflects the differences amongst pedestrians in perception of expected availability of each street into the evaluation function. The detail of the evaluation functions of utility-optimization process which we used in this study is described in Section 4.

The validity of the scheduling model which uses Mixed Logit model in order to take account of the uncertainties in the time required for each trip and the different attitudes of travellers towards them was tested by Ishikawa and Fujiwara (2002).

The process which determines local movements of pedestrians uses their viewing fields in order to avoid collision against obstacles such as walls and other pedestrian. As each pedestrian P_i has their own walking speed and direction (v_i), radius of body (r_i), and current location (x_i, y_i), the paths of all other pedestrian who are involved in the visual field defined by the size of each pedestrian's information space can be calculated for several steps ahead (*Figure 2*). When collisions are expected, the next step will be modified so that they can always keep a certain distance between other pedestrians. Spatial data which defines walkable area (streets) and non-walkable area (walls and other objects) are also used in this calculation. This collision avoidance behaviour allows pedestrians to deviate from the links and to walk "freely" in the streets which are represented as areas. This process will be refined by introducing the idea of flocking and crowding.

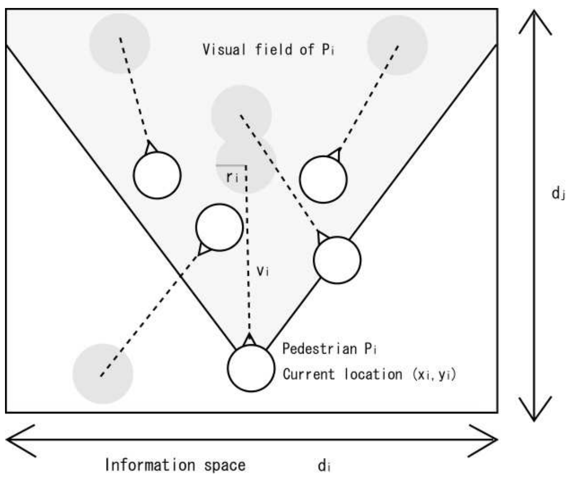

Figure 2. Local movement

4. METHODOLOGY

4.1 Data

A series of surveys on retail behavior of shoppers was undertaken at a large shopping centre in Tokyo, Japan. The centre is composed of more than 140 shops most of which are setting young women as their marketing target. The centre was selected to simplify the analysis by eliminating the influence of age and gender on behavioral patterns.

4.1.1 Street network

Figure 3 shows the street network map of the shopping centre (ground floor). The site is about 110m by 300m large and has 2 stories, and all the shops in it are charted as nodes on the network while the centre lines of the corridors are divided into links each of which has ten-meter length. Shop nodes are connected to the network by a vertical line from the nearest link. Intersections of corridors and lifts are also represented as nodes. Each node is represented as 3D coordinates (x, y, z) and has its unique ID. Each link is defined by combinations of 2 nodes. The total number of the nodes and links is respectively 326 and 364.

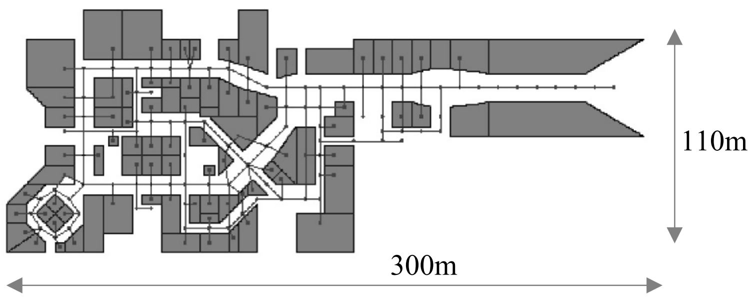

Figure 3. Street network of the shopping centre

4.1.2 Observation study

In the surveys, 18 shoppers, female graduate students who are in their twenties, were asked to shop around for 2 hours and the routes they took were tracked and recorded. For the purpose of obtaining as much detail of the movement of each shopper as possible, digital video cameras were used as main sensors of the measurement systems in this study. Every 30 seconds, the node that is closest to the shopper's location was identified from video images and recorded to duplicate the route on the network as shown in *Figure 4*.

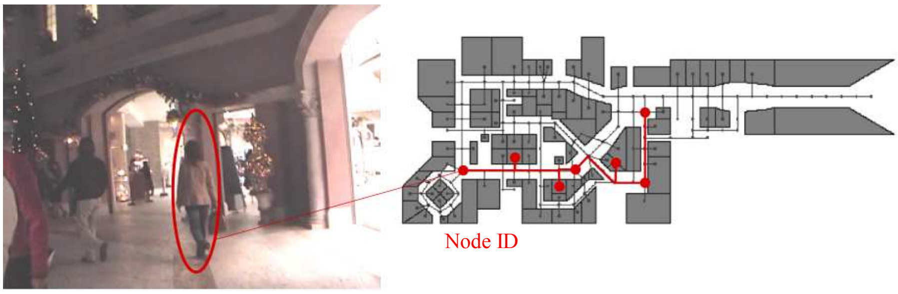

Figure 4. Tracking shoppers' movement from video image

The panel technique was used in this study in order to see the influence of both accumulation of spatial knowledge about the shopping centre itself and impression of each shop from previous experience on the routing process. The same experiments were carried on once every two weeks from 18/11/2002 to 10/01/2003 for a total of three times.

4.2 Simulation

As the scope of this paper is to develop a simplified simulation system to lay down the groundwork for future implementation of the entire model framework, we implemented a model which has a relatively simple *utility maximization* process using the shortest path algorithm as one of its evaluation criteria. It is because the route choice process, if we ignore a process of invoking new objectives of the travel by stimuli from the environment, can be regarded as a relatively simple Travelling Salesman Problem (TSP). TSP is an algorithm which, given a finite number of "places" along with the cost of travel between each pair of them, finds the cheapest way of visiting all the places and returning to the starting point. It is self-evident that the whole retail behavior or route taken by the shoppers cannot be described only by the simplified process. The comparison, however, between the computed behavior or route based on this simplified assumption and the actual route can reveal the influences of the remaining parts of the proposed framework and other influential factors.

Thus, the model used in this study assumes that each shopper goes the rounds of all scheduled shops in the shortest route. The basic structure of the simulation system developed in this study is an expansion of integrated spatial data representation systems which reconstruct pedestrian's trajectories from fragments of location data, aggregated data of traffic volumes, and common knowledge about human behaviour such as walking speed (Sekimoto, 2001, Tanaka, 2003). For the optimization process, we used Genetic algorithms and the route that each shopper took in their 2-hour travel is represented as an array of 241 IDs of the nodes as the time resolution of this simulation was 30 seconds. This array corresponds to a chromosome of which value is calculated under several evaluation criteria. This evaluation process is followed by selection, cross-over, and mutation process until the system converges on the optimum solution.

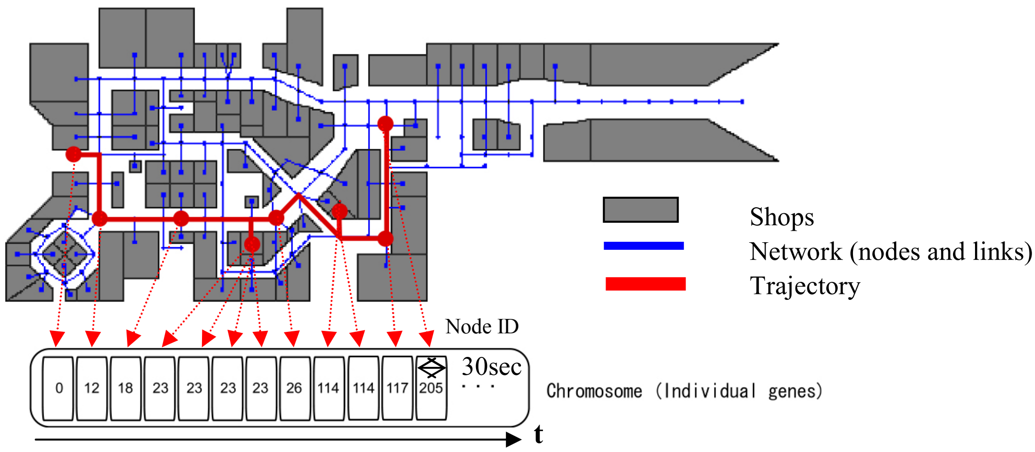

Figure 5. Representation of shopper's spatio-temporal location by Genetic Algorithms

The total value (*V*) of each chromosome is expressed as weighted linear sum of all evaluation functions and the solution and parameters which optimize *V* are searched by the system where the objective function is defined as:

$$\max V = \sum_{i=1}^{N} a_i \cdot x_i$$

(x_i : Evaluation function for criterion *i* a_i : weighting parameter for x_i)

Initialization of the GA simulation was done through random creation of chromosomes. In order to avoid unnecessary complexity of the calculation which might leads computing problems, we took account of two restrictions on the form of chromosomes. The first restriction is that the first and the last gene i.e. node ID should be the same as the pre-fixed start/goal point. The other is maximum length of travelling distance. Distance between adjacent genes is set to be less than 100m based on the average walking speed of pedestrian, 60 metres per minute. As for the evaluation criteria, we used some basic knowledge about pedestrian movement which had been identified from the observation studies:

- Physical restrictions of movements
  - ✓ walking speed 60 metres per minute
  - ✓ rotation angle less than 150 degree
  - ✓ limited vertical movements The cost of vertical movements is more than that of horizontal movements
- ID of nodes which were scheduled to visit
  Chromosomes which contains these ID will be highly evaluated
- ID of the nodes of Origin and Destination of the travel
- Travel distances (the shortest-path model)
  Standardized value of travel distance of each chromosome is derived from division of the sum of distance between all adjacent genes by the maximum length of trajectories i.e. the value of the maximum walkingspeed multiplied by the time of travel, 2 hours.

The GA process was iterated for 1000 generations and was applied to 100 chromosomes in each. In addition to crossover and mutation, we used Simulated Annealing method and Hill Climbing method for modification of chromosomes to create a new generation of them.

5. RESULTS AND DISCUSSION

Figure 6, *Figure 7*, and *Figure 8* show the results of test simulations, which used only small part of the network (38 nodes and 40 links) to construct 10-minute retail movement. The fragment of observed route that contains 2 scheduled destination points was used for a comparison with the estimated route by the shortest path model. The estimated route in *Figure 6* meets the all requirements stated in the previous chapter and its travel distance is less than that of the real trajectory. *Figure 7* shows the trajectory which did not take account of restriction on distance. The estimated route is much longer than the real one while it satisfies other requirements. On the other hand, if we tighten the restriction on distance, the estimated route tends to stay around the OD points as shown in *Figure 8*. This type of test was repeated until we found the values of parameters of each evaluation function which make the total value of the estimated route approximate that of real route. The tentative parameters' values are; 5 for walking speed, 2 for a restriction on distance, 1 for places to visit, 2 for the prefixed OD points, 0.5 for a restriction on vertical movement. We also analyzed the similarity of the 2 trajectories by calculating the distance between a pair of genes which has same t in order to test the accuracy of the parameter setting and the model.

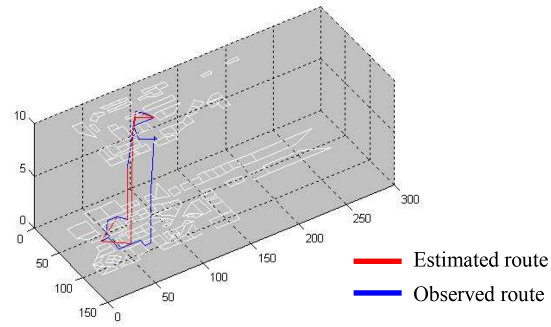

Figure 6. Test simulation

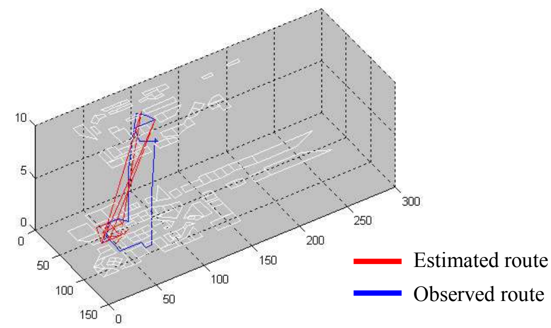

Figure 7. Test simulation without restriction on distance

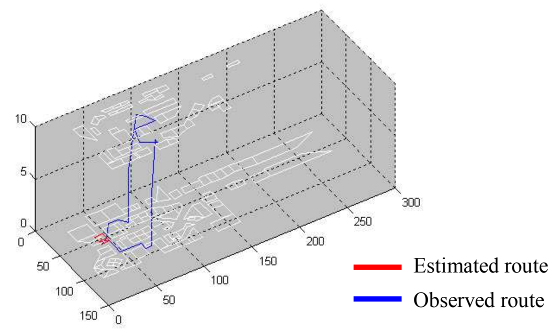

Figure 8. Test simulation with severe restriction on distance

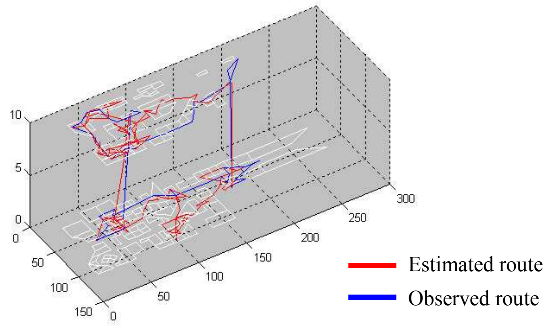

Figure 9. Simulation for whole network

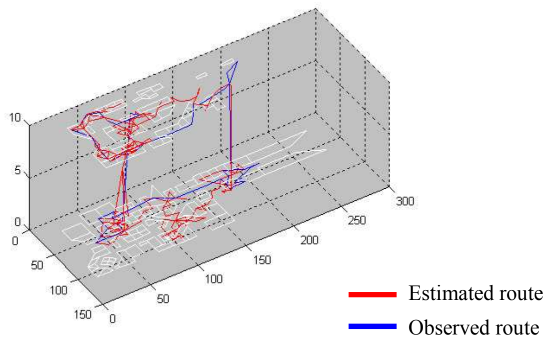

Figure 10. Simulation for whole network (the second time)

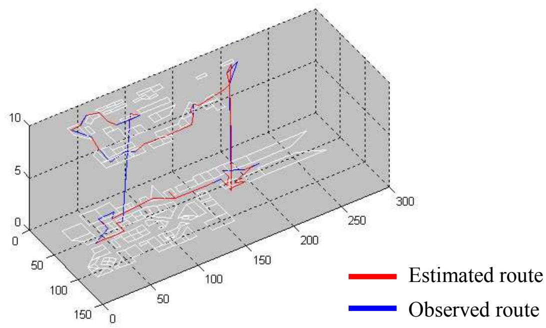

Figure 11. Simulation with true trajectory as an initial value

Figure 9, *Figure 10*, and *Figure 11* show the results of the simulations applied for the whole network. *Figure 9* illustrates that the model using *shortest path algorithm* fits relatively well to the outline of real trajectories but can not assimilate to its details. This suggests that shopper's spatial behaviour might consist of a number of segments in which shoppers follow the simplified principle of minimizing route length to visit all destinations at the route choice level. The discontinuity and deviation might be caused by local movement or information gathering processes which were proposed in section 3. It seemed also reasonable to think that the initialization process based on random walk could produce wiggle lines so as to prevent the route from remaining still and stay at the same place for a while. *Figure 10* shows the result of another simulation under the same conditions as *Figure 9*. The sum of the distance between correspondent genes in each simulation was 52.4 metres and 68.8 metres respectively. This sort of variation derives from the characteristic of GA as a quasi-optimization method. In order to avoid this problem, it might be necessary to apply statistical standardization to estimated routes. At last, we can see 2 similar routes in *Figure 11*. The total distance between the two is only 1 metre. In this simulation, the observed route was used as one of the initial chromosomes. The fact that the result is similar to the real trajectory suggests that the evaluation functions and parameters used in the model are valid to reconstruct rough routes of retail movement.

6. CONCLUSION

This study proposed a generic model of pedestrian spatial behaviour and examined the applicable scope of *shortest-path algorithm* in the model. We implemented a simulation model using GA for computational emulating of retail movements of shoppers in a shopping centre, which includes *shortest path* as one of its evaluation functions. The results suggest that the simulation model can reconstruct approximate trajectory. Further research is directed toward analysis on influential factors to implement other process of the model framework as well as improvement of validation methods.

7. REFERENCES

Desyllas, J., Duxbury, E., Ward, J., Smith, A., 2003, "Pedestrian Demand Modelling of Large CITIES: AN Applied Example from London", *Proceedings of the 8th International Conference on Computers in Urban Planning and Urban Management*, *Sendai, Japan*

Gemzøe, L., 2001, "Are Pedestrians Invisible in the Planning Process? Copenhagen as a Case Study", *Proceedings of Australia: Walking the 21st Century; An International Walking Conference*

Greater London Authority, 2004, *The London Plan; Spatial Development Strategy for Greater London 2004*, London, A copy of this document is available as of April 26 2004 at the following website. http://www.london.gov.uk/mayor/strategies/sds/index.jsp

Hoogendoorn, S.P., And P.H.L. Bovy, 2002, "Normative Pedestrian Behaviour Theory and Modelling", *Proceedings of the 15th International Symposium on Transportation and Traffic Theory, Adelaide, Australia*, p. 219-245.

Hoogendoorn, S.P., 2003, "Pedestrian travel behavior modeling", *Proceedings of the 10th International Conference on Travel Behavior Research*, *Lucerne, Switzerland*

Hughes, R., 2002, "A continuum theory for the flow of pedestrians", *Transportation Research Part B: Methodological*, Vol. 36(6), p. 507-535.

Ishibashi, K., Kumata, Y., Saito, S., 1998, "Forecasting the sales of shopping sites in the central commercial district using the disaggregate Markov Chain model: its application to evaluate the size and location pattern of shopping sites by the consumer's expenditure at kitakyusyu-city", *Journal of the City Planning Institute of Japan*, vol.33, 349-354. (Japanese)

Ishikawa, N., Fujiwara, A., Sugie, Y., Lee, B., 2002, "Scheduling Behavior Models of Long-distance Intercity Travel Considering Uncertainties", *Proceedings of Infrastructure Planning, No.26, Iwate University, Japan*. (Japanese)

Kitazawa, K., Shibasaki, R., Tanaka, H, 2003, "Study for Agent-based Modeling of Migration Behavior of Shoppers", *Proceedings of the 8th International Conference on Computers in Urban Planning and Urban Management, Sendai, Japan*

Livingstone, K., 2001, *The Mayors Transport Strategy, Greater London Authority*, London, A copy of this document is available as of April 26 2004 at the following website. http://www.london.gov.uk/mayor/strategies/transport/trans_strat.jsp

Centre Partnership and London Borough of Ealing, Regeneration & Major Project Unit, 2002, *EALING CENTRE –A Strategy For Sustainable Improvement 2002-2012*, London, A copy of this document is available as of April 26 2004 at the following website. http://www.ealing.gov.uk/council/strategies+and+policies/town+centre+plans/

Nagel, K., Beckman, R. J., Barrett, C. L., 1998, "TRANSIMS for transportation planning", *Los Alamos National Laboratory Unclassified Report*, LA-UR 98-4389 LANL, A copy of this document is available as of April 26 2004 at the following website. http://www-transims.tsasa.lanl.gov/

The Royal Borough of Kensington and Chelsea, 1997, *Kensington High Street Action Plan*, London, A copy of this document is available as of April 26 2004 at the website below. http://www.rbkc.gov.uk/EnvironmentalServices/general/hskplans.asp

Schelhorn, T., O'Sullivan, D., Haklay, M., Thurstaion-Goodwin, M., 1999, "STREET: An Agent-based Pedestrian Model", *Centre for Advanced Spatial Analysis working paper* series 9. A copy of this document is available as of April 26 2004 at the website below. http://www.casa.ucl.ac.uk/working_papers.htm

Sekimoto, Y., 2001, *A Basic Study on the Reconstruction of Dynamical Change of Real World Integrating Many Kinds of Data*, Doctoral thesis submitted to the Department of Civil Engineering, University of Tokyo (in Japanese)

Tanaka, H, 2003, *A Basic analysis for modelling migration behaviour in shopping area*, Masters' thesis submitted to the Department of Civil Engineering, University of Tokyo

Thompson, P.A., 1994, *Developing New Techniques for Modelling Crowd Dynamics*, PhD thesis submitted to University of Edinburgh