Short Paper

Ref 135

Evaluated Model of Pedestrian Movement Based on Space Syntax, Performance Measures and Artificial Neural Nets

F L Zampieri
UFRGS, PROPUR, Porto Alegre, Brazil
fllz@yahoo.com.br

D Rigatti
UFRGS, PROPUR, Porto Alegre, Brazil
driga2000@yahoo.com.br

C Ugalde
UFRGS, PROPUR, Porto Alegre, Brazil
claudiougalde@yahoo.com.br

Keywords
pedestrian movement; pedestrian modelling; space syntax; level-of-service; artificial neural nets

Abstract

To understand the generating causes of pedestrian movement is very important for urban planning tasks, because it is possible to infer if attitudes taken in the conception and maintenance of the spaces are in fact contributing to the social dynamics. However, to determine pedestrian flows is a difficult task due to the complexity of people movement. A way to outline this problem is through the creation of models which associate the attributes and their relationships directly to the studied phenomena. The model used here uses two kinds of variables: the configurational ones obtained through the axial map of the city where the study area is located, and the performance measures obtained through the physical evaluation of the attributes of the sidewalks of the studied area. The output of the model is the mean pedestrian rate of the area. The Syntax Space theory is useful for understanding the phenomenon because of the way it deals with space interaction. However, though it is able in predicting part of the movement, we do not find significant correlations when the measure of intelligibility is low. Pedestrian flows are a complex phenomenon and, per se, cannot be understood through linear relationships among any couple of variables, being them spatial or not. In this paper it is argued that the space syntax theory and measures explain the pedestrian movement as a phenomenon emerged from society, but the linear approach is not capable of explaining their relationships. The presented model uses Artificial Neural Nets (ANN), a parallel processing tool with the capacity of working through examples, learning, generalizing and abstracting the variables information and their connections. The implementation of these kinds of models evolves from 'black boxes' to models that can be 'disassembled' and evaluated inside of its logical structure. The ANN uses two groups of data: one for training nets and the other one for validating the network. Thus, the performance of the ANN can be tested with unknown data. The results produced so far have shown that ANN can learn the main features of the data sets with an accuracy of more than 90% of correlation coefficient and with an average error smaller than 0.02. It must be said that the research work targets to spread the samples from different configurational realities and expand the data bank on measured movement in order to improve the accuracy of the model, such as being done lately.

1. Introduction

Walking is the basic form of the human being's movement. However, people don't simply overcome distances. They move within cities with different objectives, such as to do shopping, to socialize, to go to some particular place or only to wander. Everyone moves in his own way, with their unique mental and physics distinctive peculiarities and motivations. However, human behaviour is based on a number of shared decision-making parameters and spatial accessibility restrictions imposed by the environment where it occurs.

"Walking" is doubtless a transportation mean because as any other modal, it depends on intrinsic (of the user) and extrinsic (of the environment) factors. However, if we do not consider it in all its dimensions we are simplifying the reality by reducing the way we understand the pedestrians behaviour, disregarding important attributes like: ambiance; socialization elements; health and rest; and aesthetic.

The comprehension of how pedestrian flows happen in urban areas is impossible to detect through phenomenon observation because the movement is a complex phenomenon originated from how society construct and manages the space. Space syntax theory is useful once it decomposes the urban grid, where the movement occurs, into lines making it possible to understand the relations among spaces. Consequently, space syntax theory is capable to explaining the pedestrian flow because it is considered as a phenomenon which is originated by people's displacement.

When linear methods are used together with space syntax theory it is not possible to explain all relationships among the variables. As explained by Mak et al. 1996 and Olaru et al. 2001, linear methods tend to present poorer correlations when compared to ANN and other non-linear techniques. Non-linear systems are capable to recognize the complexity that emerges from the urban grid pattern, not as laws or strict rules but as a range of possibilities. Studies involving complexity assume there are no rules, but patterns that modify themselves through the evolving society and space.

The urban pedestrian movement influences the growth and the form of the cities. In order to improve movement and to reduce negative impacts, a great number of researchers are using modelling techniques for representing and understanding the phenomena.

On this paper we propose a way of modelling pedestrian movement, using a model in which space syntax measures are combined to other qualified environmental parameters in order to identify, understand and pounder the spatial attributes that contribute to enhance pedestrian movement.

2. Pedestrian Models

There are several kinds of descriptive representations for pedestrian movement and the environment where they occur. By building a model of that behaviour, we seek to understand the characteristics related to the pedestrian behaviours. The difficulty on the pedestrian flow analysis starts from the lack of tools for modelling and simulating the variation in the movement rates, in terms of changes in the urban environment to the validation of such variation through real data.

A large number of models have been developed to provide explanations for the pedestrian movement phenomena. A wide range of theories where used for that, such as: queuing models (Hoogendoorn and Bovy, 2004; Løvås, 1994), transition matrix model (Helbing et al. 2001; Kurose and Hagishima, 1995), stochastic model (Ashford, 1976), route choice model (Hoogendoorn and Bovy; 2004). They are all, at least partially, similar to each other and use to describe pedestrian behaviour in critical situations like in fire exits, emergency evacuations, among others. However, none of these models take into account the effects of the self-organization of pedestrian groups. Another approach proposes the analogy between pedestrian movement and the behaviour of some gasses and liquids (Helbing et al. 2001). However, in order to be useful to modelling pedestrian movement, specific features are highlighted like pedestrian intentions and interactions.

The above mentioned models have their applications limited due to their complicated formulations which make it difficult to use in real case scenarios.

In order to understand how, and which environmental variables influence the pedestrian movement level-of-service (LOS) models are used to evaluate the quality of the sidewalks and the perception of both safety and comfort of pedestrians. The relationship among pedestrians their environment and the sidewalks conditions are essential to understand pedestrian movement, which is much more complex when compared to the movement of vehicles.

Differently from the above mentioned models, configurational models like those propose by Space Syntax assume that cities are mechanisms that generate a potential field of social encounters and co-presence (Hillier et al., 1993). In other words, it is not the effect of the city but the social effect that structures the potential of the pedestrian movement and not merely the movement to building entrances as usually considered in transport studies.

The method used in this paper targets to make it possible to understand the urban characteristics initially from studying the spatial structure of the whole city (global level) and the sidewalk itself (local level) with the support of an urban analysis model based on artificial neural nets (ANN). Therefore, this paper deals with two different methodologies: one from the space syntax theory and the other from the level-of-service studies (LOS). Variables and their associations can be explored, creating as many scenery types as necessary to verify, through simulation, the urban environment performance. The space syntax, through the syntactic measures of the axial map, and the level-of-service, through the performance measures are used in this model as input variables and the pedestrian flow is the output of the model processed with ANN techniques.

The challenge is to create a model that works within different system parts without a specific rule, since it is based on emergency patterns that drive complex systems. With the ANN there is no need to defining the way the relations among the variables happen making it possible to explore the degree in which spatial design and sidewalks performance are related to the pedestrian flow. ANN allows building several sceneries with limited variables, without a defined algorithm.

3. Artificial Neural Nets

The artificial neural networks (ANN) are a computation type based on the human brain's neural nets that processes information in parallel – meaning all the input data is processed at the same time (Haykin, 2001). Initially, the ANN trains part of the input data, which is called the learning phase or training. In a second phase the test is accomplished with the part of the data that wasn't initially trained examining if the learned characteristics can be applied in the non processed data, unknown by the ANN. High correlations and low errors mean that the phenomenon's characteristics have been learned. The capacity of learning and generalizing information is what makes the neural nets such an interesting tool for solving complex urban problems such as pedestrian movement.

Three different methods are very important in the ANN approach: relative importance of the variables; sensibility test and evaluation of the variables co-dependence. Garson (1991) developed a method of analysis that can determine the importance of the variables for the model through the weights of the neural connections. The sensibility tests are made in the models in order to determine in which extent each input variable influences the output (Olden et al. 2004; Lek et al. 1996). Finally, the evaluation of the co-dependence among the variables is developed through an electronic spreadsheet with the inputs of data model.

4. Application Of The Model

The broad study case was taken from the central areas of three different Southern Brazilian cities, including more than five hundred sidewalks. This paper will show only the results for Santa Maria City, comprising 71 sidewalks. It focusing two main aspects concerning the sidewalks sample: those should represent the most different used for pedestrian movement and the built environment surroundings should display a consistent diversity of land-uses which characterizes centrality.

Figure1
Location of the study case Santa Maria City

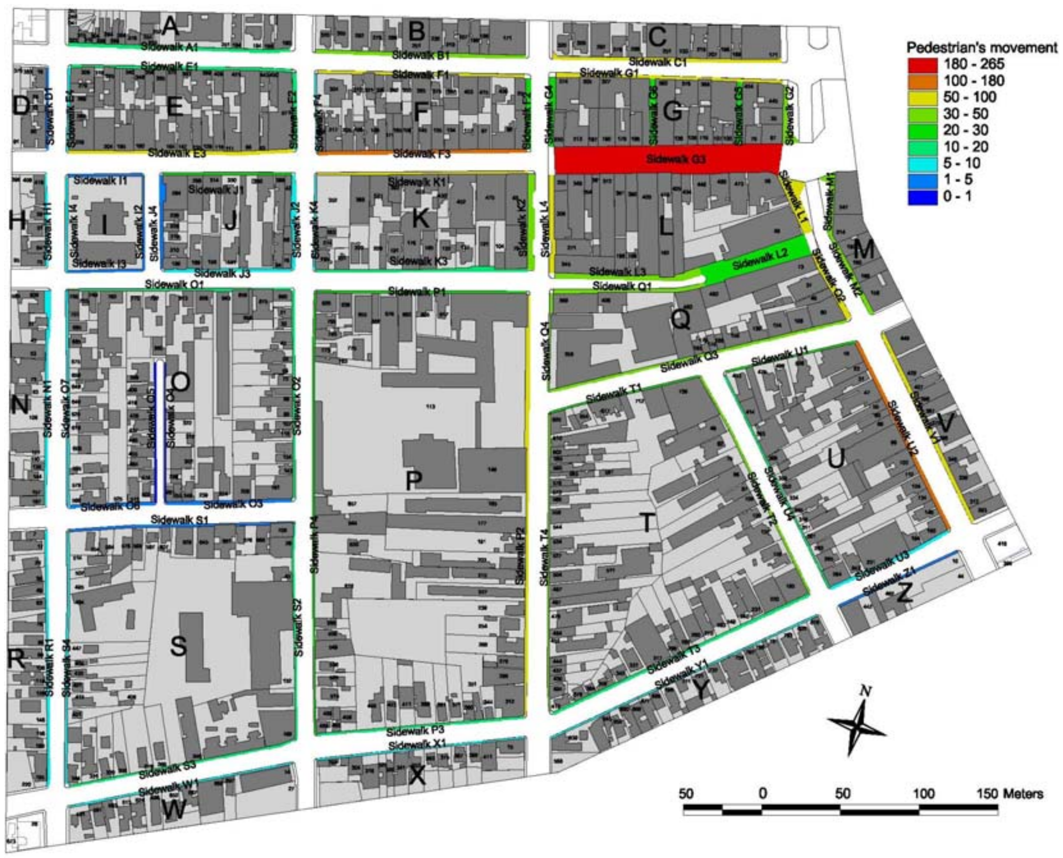

Figure 2
Map of the pedestrians' movement

The pedestrian movement was obtained counting the moving people in each sidewalk in five different periods a day and for five weekdays for each sidewalk (proposed by Hillier et al. 1993), resulting in a total of 50 measurements for each sidewalk from which the mean movement rates were obtained (Figure 2).

The syntactic variables obtained from the axial map for the entire city are: global integration (RN); local integration (R3); connectivity; control and depth. Other measures related to attractors are: constitutions, residential attractors, commercial attractors (Figure 3), service attractors and other attractors.

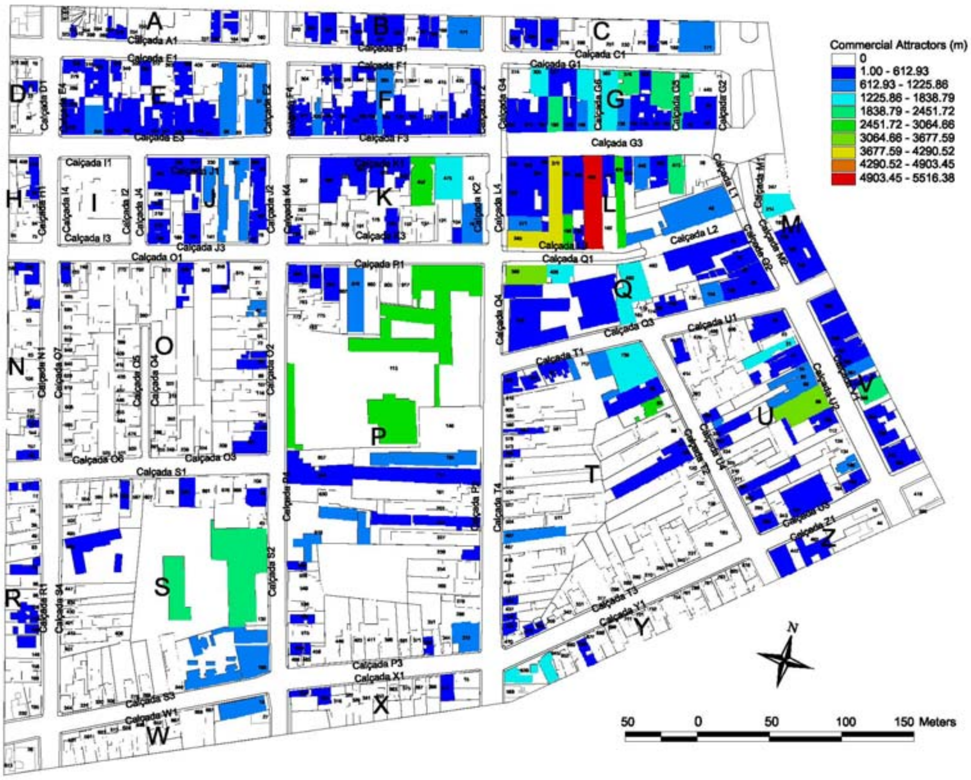

Figure 3
Map of the 'commercial attractors' in meters

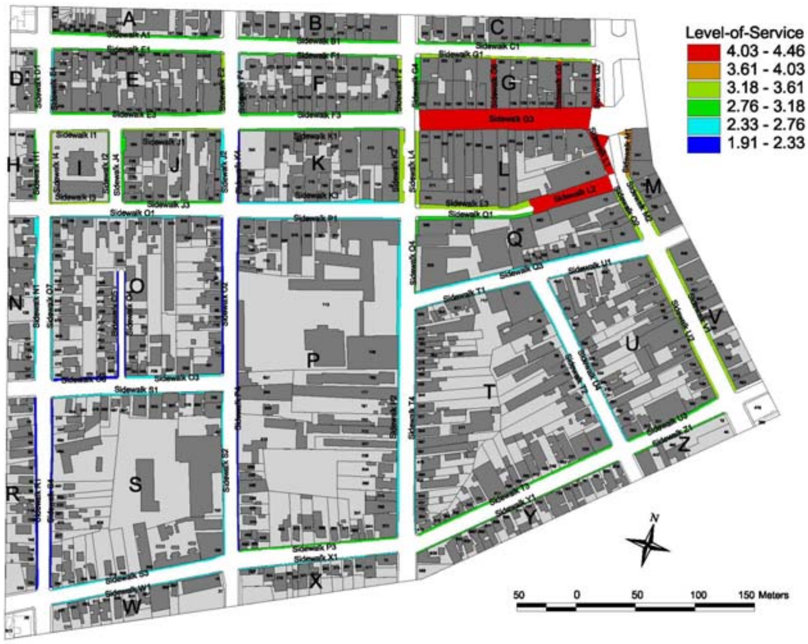

Figure 4
Map of the 'Level-of-service'

LOS in this paper is adapted from the Highway Capacity Manual (2000) and Khisty (1994) methodologies in both quantitative (width and length) and qualitative (attractiveness, comfort, maintenance, safety and public safety) attributes of the sidewalks, resulting in the evaluation of the quality of the sidewalks (performance measures), varying from "A" (best quality) to "F" (the worst).

5. Results

The model processed 75% of the data set in a neural nets software as the training set and the 25% remaining data as the unknown testing data. The results are reproduced in an electronic spreadsheet based on the sigmoid logistics function of the ANN where the number of pedestrians is simulated and compared with the movement rates found in the actual area. The results showed an R2 of 0.9437 for the training set (RMS of 0.0014689), and a R2 of 0.9611 for the unknown data (RMS of 0.0024692) allowing to infer that the phenomenon can be explained through the used variables. Despite the fact that high correlations indicate the effectiveness of the model it is necessary to verify how the variables work in its internal structure through the analysis of the importance of the variables (Figure 5), a sensibility test (Figure 6) and an evaluation of the co-dependence among the variables (Figure 7).

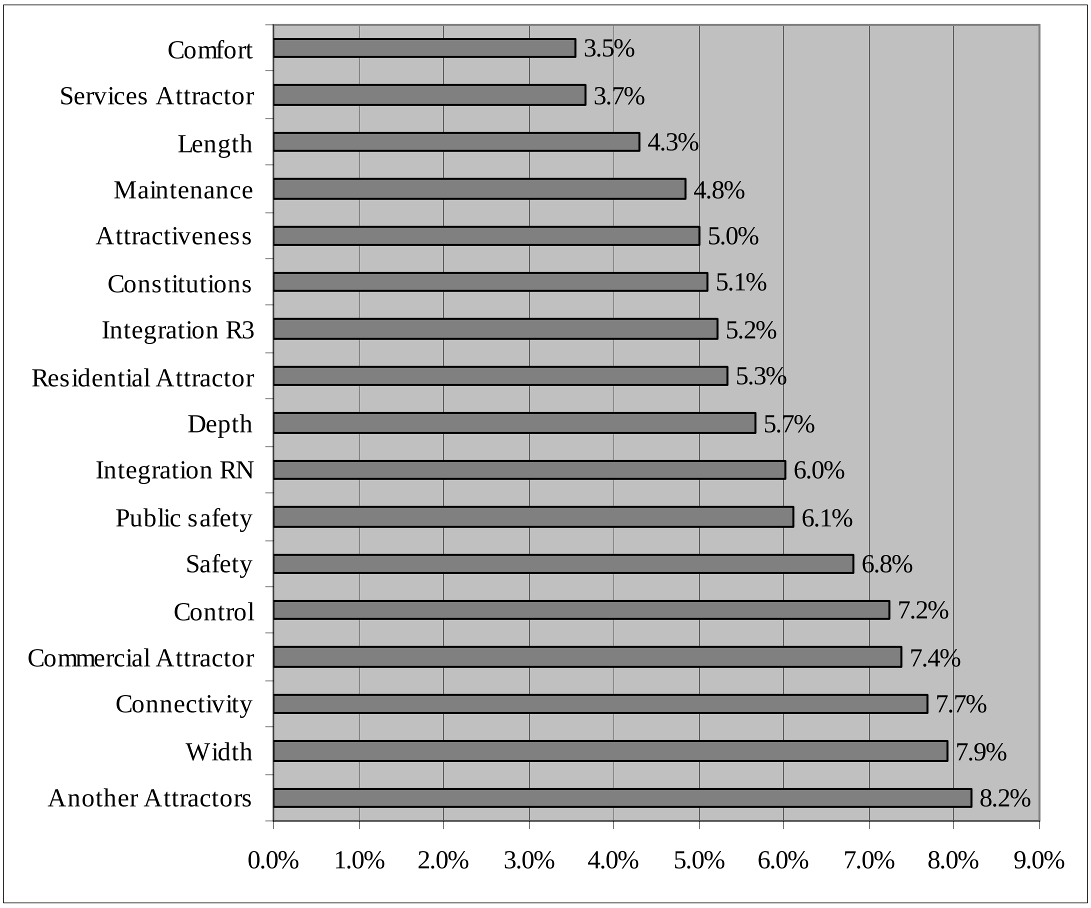

Figure5
Importance of the variables (Garson, 1991)

By analyzing the figures 5 and 6 we can observe how the inputs 'commercial attractors' (CA) and 'constitutions' (CT) work. Even if the CA is one of the Garson's items of greatest importance and displays a rising curve in the sensibility test the CT, which has a low importance for Garson, shows a great potential for pedestrian flows which enables us to corroborate syntax theory hypothesis once more. The ANN learns how the inputs of the system work and the sum of effects of the variables also can increase the simulated flows. The figure 7 shows how the model responds to the variables variation. Both the CT and the CA can raise the pedestrian flow rates but if they are taken together (Figure 7, scenario 6) and considering the values of 'global integration', 'connectivity' and 'public safety' in the maximum, the rate remains almost the same. On the other hand, if the variables like 'integration R3', 'control' and 'depth' are considered in the maximum, the model radically decrease the movement rates (Figure 7, scenario 10).

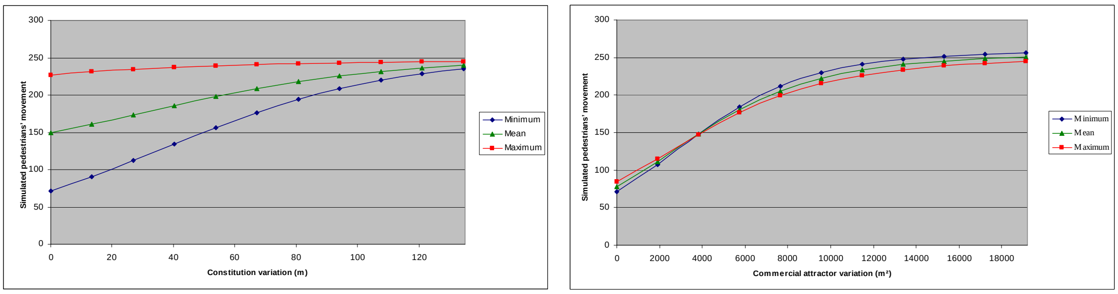

Figure 6
Sensitivity test of the 'commercial attractors' and 'constitutions'

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
1	**Model normalized between 0 and 100%**								Enter with a value between 0 (the minimum of the model) and 100 (maximum of the model)											
2	**Inputs**	**Integration RN**	**Integration R3**	**Connectivity**	**Control**	**Depth**	**Constitutions**	**Residential Attractor**	**Commercial Attractor**	**Services Attractor**	**Another Attractors**	**Width**	**Length**	**Attractiveness**	**Comfort**	**Maintenance**	**Safety**	**Public safety**	**Output**	**Pedestrian Movement**
4		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
5	Scenario 1	0	0	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0		256.168
6																				
7		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
8	Scenario 2	0	0	0	0	0	100	0	0	0	0	0	0	0	0	0	0	0		235.216
9																				
10		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
11	Scenario 3	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0		97.0315
12																				
13		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
14	Scenario 4	0	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	0		71.4737
15																				
16		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
17	Scenario 5	100	0	100	0	0	0	0	0		0	0	0	0	0	0	0	0		151.209
18																				
19		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
20	Scenario 6	100	0	100	0	0	0	0	0	0	0	0	0	0	0	0	0	100		243.265
21																				
22		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
23	Scenario 7	100	0	100	0	0	0	0	100	0	0	0	0	0	0	0	0	0		254.024
24																				
25		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
26	Scenario 8	0	0	0	0	0	0	50	0	25	0	0	0	0	0	0	0	0		99.0352
27																				
28		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
29	Scenario 9	100	0	100	0	0	0	50		25	0	0	0	0	0	0	0	0		159.087
30																				
31		OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK	OK		
32	Scenario 10	0	100	0	100	100	0	0	0	0	0	0	0	0	0	0	0	0		3.71589
33	**Variation of data entrance.**																			
34	MIN	0.70831	1.56692	1	0.09091	1	0	0	0	0	0	1.4	45.4	1.5	1.83333	2.5	1	1		0.24
35	MAX	0.83139	5.24404	20	4.49026	5	134.5	21442.5	19128	8610.22	13889.8	22	330.9	4.33333	5	4.5	5	4.66667		261.8

Figure 7
Evaluation of co-dependence variables

6. Conclusion

The benefit of any model is to allow the visualization of the concepts used and the existing attributes through the representation of the reality. A correlation coefficient higher than 90%, like in this case, can be considered statistically strong, although the results can not be generalized for others cities. The high correlation found indicates that this methodology can also be used on similar researches.

The spatial syntax measures are very useful to the model because they allow us to understand the influence of the urban layout on the natural movement. The model showed a strong correlation between 'global integration' and 'connectivity' when used together: thus, if the urban system has a strong intelligibility, which is not the case, (R2 = 0.16) the probability is that just the 'global integration' can determinate the movement rates. The attractors act as amplifiers of the grid characteristics increasing the pedestrian flow when they are linked to the residential and service attractors (Figure 7, scenario 8 and 9). The analysis method sought by this paper was not only the prediction of pedestrian flow but to propose a tool for understanding and discussing movement rates according to a large number of variables. Although limited in case studies, the results are quite promising as a model for prediction/evaluation of the pedestrians flows. Even if models can't predict exactly what will happen under a certain social and spatial decision circumstance, they make it possible to relate the events to their causal sources and contribute to sharpen decision making in urban planning environments.

References

Ashford, N, M O'Leary, and P. D. McGinity. "Stochastic Modelling of Passenger and Baggage Flows through an Airport Terminal." *Traffic Engineering and Control* 17 (1976): 207-10.

Garson, G. David. "Interpreting Neural-Network Connection Weights." *AI Expert* (1991): 47-51.

Helbing, Dirk, Péter Molnár, Illés J. Farkas, and Kai Bolay. "Self-Organizing Pedestrian Movement." *Environment and Planning B: Planning & Design* 28 (2001): 361-83.

Highway Capacity Manual. Washington, DC: TRB Publication, 2000.

Hillier, B., A. Penn, J. Hanson, T. Grajewski, and J. Xu. "Natural Movement: Or, Configuration and Attraction in Urban Pedestrian Movement." *Environment and Planning B: Planning and Design* 20 (1993): 29-66.

Hoogendoorn, S. P., and P. H. L. Bovy. "Pedestrian Route-Choice and Activity Scheduling Theory and Models." *Transportation Research Part B*, no. 38 (2004): 169-90.

Khisty, C. Jotin. "Evaluation of Pedestrian Facilities: Beyond the Level of Service Concept." *Transportation Research Record*, no. 1438 (1994): 45-50.

Kurose, Shigeyuki, and Satoshi Hagishima. "A Method for Identifying Accessibility Properties of Pedestrian Shopping Networks." *Journal of Retailing and Consumer Services* 2, no. 2 (1995): 111-18.

Løvås, Gunnar G. "Modeling and Simulation of Pedestrian Traffic Flow." *Transportation Research part B: Methodological* 28, no. 6 (1994): 429-43.

Mak, Brenda, Tung Bui, and Robert Blanning. "Aggregating and Updating Experts' Knowledge: An Experimental Evaluation of Five Classification Techniques." *Expert Systems With Applications, Vol. 10, No. 2, pp. 233-241* 10, no. 2 (1996): 233-41.

Olaru, Doina, and Raluca Raicu. "Accessibility Measures of Transport Networks. Comparison of Neural Networks and 'Classical' Approaches." In *The 23rd Conference of Australian Institutions in Transport Research*. Melbourne, Australia, 2001.

Olden, Julian D., Michael K. Joy, and Russell G. Death. "An Accurate Comparison of Methods for Quantifying Variable Importance in Artificial Neural Networks Using Simulated Data." *Ecological Modelling* 178, no. 3-4 (2004): 389-97.