Procedia Computer Science
Volume 51, 2015, Pages 2943–2947
ICCS 2015 International Conference On Computational Science

Agent Based Simulations for the Estimation of Sustainability Indicators

Ander Pijoan, Cruz E. Borges, Iraia Oribe-Garcia, Cristina Martin, and Ainhoa Alonso-Vicario*

Deusto Institute of Technology – DeustoTech Energy,
University of Deusto, Avda. Universidades 24, 48007 - Bilbao, Spain
{ander.pijoan,cruz.borges,iraia.oribe,cristina.andonegui,ainhoa.alonso}@deusto.es

Abstract
We present a methodology to improve the estimation of several Sustainability Indicators based on the measurement of walking distance to infrastructures combining Agent Based Simulation with Volunteer Geographic Information. Joining these two forces we construct a more realistic and accurate distribution of the infrastructures based on knowledge created by citizens and their perceptions instead of official data sources. A Situated Multi-Agent System is in charge of simulating not only the functional disparity and sociodemographic characteristics of the population but also the geographic reality in a dynamic way. Namely, the system will analyze different geographic barriers for each collective bringing new possibilities to improve the assessment of the needs of the population for a more sustainable development of the city. In this article we will describe the methodology to carry on several sustainability indicator measurements and present the results of the proposed methodology applied to several municipalities.

Keywords: Sustainability Indicators, Accessibility Studies, Agent Based Simulation, Volunteer Geographic Information

1 Introduction

Sustainability indicators emerge as a great set of signals with which estimate progress, analyze trends and help future decisions within this new development model. Accessibility is defined as the proximity of citizens to basic services and seeks to assess the needs of the population, regardless of their functional disparity and sociodemographic characteristics. In this sense, the consensus to define a infrastructure as accessible is that it can be reached in less than five

*This research has been carried out in the following projects: a) GREEN TRAVELLIN (Era NET Transport 6/12/IN/2014/195 label) partially founded by Diputación Foral de Bizkaia through the Plan de Promocin de la Innovacin (6/12/IN/2014/195). b) SEPARA and BIZKAISENSE projects also founded by partially founded Diputación Foral de Bizkaia through the BizkaiLAB call. c) CIUDADES AMIGABLES partially funded by the Portugalete Town Hall.

Selection and peer-review under responsibility of the Scientific Programme Committee of ICCS 2015

doi:10.1016/j.procs.2015.05.488

minute walking [6]. That is, taking that the speed of a pedestrian is approximately $5\,\mathrm{km\,h^{-1}}$, the distance to the infrastructure must not exceed 450 m. Near presence of basic services indicates that the urban fabric is particularly suitable for habitation, provides resources to make social life on the streets and avoids unnecessary travel by motorized vehicle. In addition, there already exist some studies proving having basic amenities near, is highly related with a more active and healthier lifestyle [4].

Accessibility is often calculated without the appropriate level of detail or using data sources that do not represent the feelings and point of view of the citizens. In this article we will present two methodologies to improve this situation using Volunteer Geographic Information (VGI) and Open Data. We will take advantages of the Geosimulation concept to model how the citizens move through the city. The use of this approach is not new; for example it has been used to model long term load forecasting [5] or traffic congestion [1].

Therefore, the rest of the paper is structured as follows. In Section 2 we will explain the data and methods used, in Section 3 we will show and discuss the results obtained, and finally, in Section 4 we will draw the conclusions of the paper and outline the future work.

2 GIS Data and Routing

VGI provides a whole new virtual layer where spatial data can be accessed and understood by any computational system. OpenStreetMap (OSM) is the most known VGI source; its user and developer community have generated a large database of comprehensive geographic information and tools to edit, view or analyze it in a easy and friendly way. Depending on the area, mappers have reached different levels of detail in the mapped features. Thus, it is possible to find from highly precise mapped areas to zones with only basic elements such as roads [2].

2.1 Existing Methodologies

Taking a look at different sustainability reports, methodologies currently used by local authorities are mainly based on creating areas of influence around basic services [3]. Most optimistic methodologies, are based on the spatial intersection between districts and the 450 m expanded location. Reports using this technique, assume that the population of all districts and sections that intersect with the expanded location, have walking access to the amenity. Doing so, the certainty of the results varies greatly depending on the division of districts and sections: in cities, districts and sections only cover a small number of blocks and the intersection of districts is very similar to the result of taking only the buildings within the radius of 450 m. However, small villages tend to be divided into a very few districts of which one contains the entire rustic land. For example, Figure 1 shows one of these districts used to cover the entire rural land. In those cases, this methodology presents unreliable results, determining that citizens who live kilometers away from services have walking access to them.

As an improvement to the district based estimations, there are more accurate methodologies which only take buildings included within the area of influence of the basic services or dividing a city into a grid of cells. Although these present substantial improvements in the results, they do not take into account geographical and architectural barriers such as rivers, highways or other which require a detour to cross over. That is the reason why these estimations normally lack adequate accuracy and/or can be improved using more appropriate techniques. In addition, depending on the age or functional disparity of the citizens, 450 m do not always correspond to 5 minute walking. In view of this, we present in this article two methodologies, a first approach using only GIS and a second one enhancing a GIS by geosimulations.

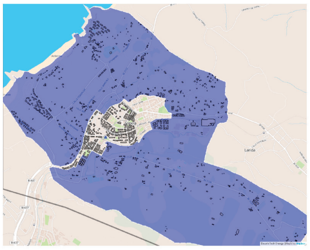

Figure 1: Rural district

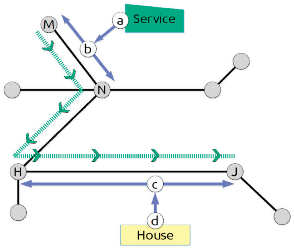

Figure 2: Routing example

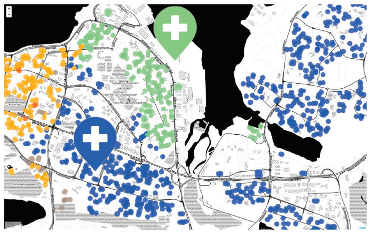

Figure 3: Nearest hospitals to houses

2.2 Door to Door Routing Methodology

Having a rich dataset, it can be possible to improve substantially the previous methodology by using routing algorithms. As we are interested in measuring the time needed to travel a route, it is necessary not only to set the starting and finishing point but also the velocity of the traveler. In this case we have fixed it at $1.42\,\mathrm{m\,s^{-1}}$ (see Section 1).

Normally, buildings have an entry point connected to the road network but if they do not, it is assumed that the nearest point to a road will be their entrance (`a`). Once all entrances have been found, we calculate the distance from each entrance to their nearest road point `d(a,b)`. Note that `b` does not have be a node of the routing graph so the algorithm cannot start from this point. The next step is to calculate the distance from `b` to the road start `d(b,M)` and end `d(b,N)` points, which really are nodes of the graph. The same is done with buildings, calculating their entrance (`d`) and distances to nearest graph nodes (`d(d,c)`,`d(c,H)`,`d(c,J)`). Figure 2 shows a diagram to ease the understanding of the procedure.

At this point a multiple destination shortest path Dijkstra algorithm is launched from the nodes of the service `M,N`. This algorithm calculates the cost of traveling from these two nodes to all building nodes (in this case `H,J`). It returns the cost of going from `M` to `H,J` and from `N` to `H,J`, that is, the four possible combinations. Finally, once all distances and times are known we select the route with the lowest cost for reaching the basic service.

2.3 Agent Based Door to Door Routing

Despite having reached a level of accuracy more than acceptable, calculations still need to take into consideration more social aspects. It is well known that citizens do not always take the shortest path as there are streets some collectives can not cross (blackspots, dark streets, steep slopes, steps. . .) and others they prefer going through (wide streets, with shops, pedestrian streets. . .). Furthermore, in the previous methodology we have assumed that all people can travel at $1.42\,\mathrm{m\,s^{-1}}$. This is hardly true as different collectives may struggle to keep this pace. Here is where Geosimulations comes in by allowing us to model different interest groups like youth, seniors or people with disabilities and personal preferences.

To this end, the Geosimulation emulates the urban dynamic by creating one agent per citizen. These inhabitants are spread evenly on the buildings of the municipality weighting the probability in base of the total surface of the buildings (taking into account the floor number). Each agent creates its own routing graph according to its preferences by assigning different priorities to roads and eliminating streets it can not use. Next, the shortest route which fulfills the agents preferences is calculated. Finally the agent will travel along that route according to its determined speed.

Table 1: Percentage of people that have foot access to this infrastructures in Reykjavík for every methodology.

	EDU			HEALTH			BANK		
Time (min)	**E**	**R**	**A**	**E**	**R**	**A**	**E**	**R**	**A**
$x < 5$	80.70	31.08	17.25	9.43	2.79	1.43	0.37	0.35	0.31
$5 \leq x < 10$	18.62	34.44	23.56	15.34	8.02	3.69	1.72	0.19	0.10
$10 \leq x < 15$	0.68	19.43	18.33	13.56	8.77	5.47	2.60	0.53	0.21
$15 \leq x$	0.68	34.48	59.19	75.23	89.19	94.88	97.91	99.46	99.59

3 Experimental Results

The proposed methodologies have been contrasted in the municipality of Reykjavík, capital and largest city of Iceland. We have analyzed the distances to the following infrastructures: educational centers (EDU), clinics and hospitals (HEALTH) and financial amenities (BANK).

Table 1 contains the percentage of citizens that live at less than 5, 10 and 15 minutes of the different infrastructures analyzed. Columns denoted by E represent the numerical results for the intersection methodology, denoted by R results for door 2 door outing methodology and, finally, columns denoted by A the results for agent based door 2 door routing.

As expected, the intersection methodology highly overestimates the number of citizens that have good accessibility to the infrastructures analyzed. In fact, in the case of EDU infrastructures it went from 81 % to 17 % when compared to the agent based methodology. Please note that this results are approximated (see note in Subsection 2.3). Another expected result is finding very low percentage of citizens who have access to HEALTH and BANK. In big cities those infrastructures tend to be centralized either in the periphery (HEALTH) or down town (BANK) and be supported by public transport.

Figures 4, 5 and 6 present the density plot of the distances of the citizens to EDU, HEALTH and BANK infrastructures respectively. In these plots it is represented the "amount" of citizen that lives at a certain distance to those facilities in a continuous fashion. The blue curve represent the distribution obtained from the intersection methodology, red for results of the door 2 door routing methodology and, finally, the green curve for the agent based door 2 door routing. Please note that the results of Table 1 are just the area under different sectors of the curve.

This figures show more clearly the differences between the three methodologies: The blue curve is always *peaker* near 0 and then falls faster than the other two. On the other hand, the green curve has a more flat and spread peak. Note that this is partially explained as the agent based routing have always worse velocities that the routing methodology. More importantly, the green curve is not an scaled version of the blue one. This is due to the fact that the distribution of the agents clearly affects their walking time to the amenities which is clearly reflected in those figures.

Another clear distinction comes when comparing the *shape* of the three curves between the different infrastructures. For example, the curves in Figure 4 resemble a χ^2 Distribution. On the other hand, the curves in Figure 5 and 6 have a more irregular shape. It can be easily explained taking into account that if citizens distribute themselves Gaussianly around those infrastructures. Then, the distance trivially follows a χ^2 Distribution.

Another remarkable results can be found in the vicinity of the population to hospitals. The fact that its geography has many peninsulas, bays, straits and small islands, makes direct distances not to correspond with routing distances. An illustrative example can be seen in Figure 3, showing buildings that despite having a relatively nearby hospital, the detour needed to reach them makes another hospital to be closer. The Figure 5 also reflects this results as it clearly contains three peaks.

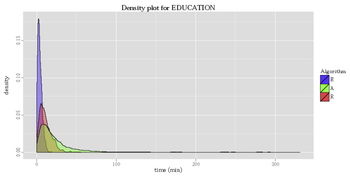

Figure 4: Density plot EDU

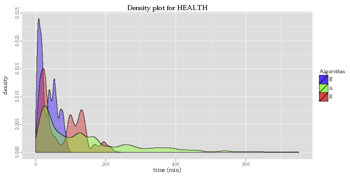

Figure 5: Density plot HEALTH

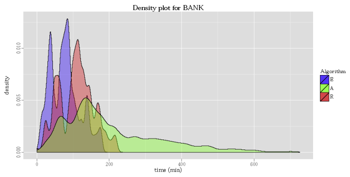

Figure 6: Density plot BANK

Finally, the curves of Figure 6 represent the typical infrastructure that is highly concentrate in the city center. In this way, citizens living down town are the only ones that are near leaving the vast majority of the population, which lives in the suburbs, far away. Please note that the three small peaks that can be seen in the blue and red curves are artifacts of the kernel used to create the density plot. Using a wider kernel those peaks would disappear.

4 Conclusions and Future Work

The work presented here is an experimental methodology for estimating better sustainability indicators. Results have proved to match in a more realistically way the humans behavior patterns and provide a better dashboard of the necessities of the citizens to cover the objectives of municipal strategic programs. Furthermore, these techniques provide a framework where to simulate future or hypothetical scenarios and analyze the effects of the location of new infrastructure in certain areas in order to assist decision-making. Despite having reached a level of accuracy more than acceptable, the methodology is still young and could refine the calculations much. On the one hand, taking full details of: presence of sidewalks, road crossings, traffic lights, barriers, walls and private walkways could provide a much more accurate estimation of the routes. On the other, having access to the census would improve the quality of the sociodemographic factors that come into play in the urban dynamic.

References

[1] Jrg Dallmeyer, AndreasD. Lattner, and IngoJ. Timm. Gis-based traffic simulation using osm. In Guido Cervone, Jessica Lin, and Nigel Waters, editors, *Data Mining for Geoinformatics*, pages 65–82. Springer New York, 2014.

[2] Mordechai Haklay. How good is volunteered geographical information? A comparative study of OpenStreetMap and Ordnance Survey datasets. *Environment and Planning B: Planning and Design*, 37(4):682–703, 2010.

[3] City of Reykjavík. European Green Capital Application, Reykjavík, Iceland. Technical report, 2012.

[4] Frank J Penedo and Jason R Dahn. Exercise and well-being: a review of mental and physical health benefits associated with physical activity. *Current opinion in psychiatry*, 18(2):189–193, 2005.

[5] A. Pijoan, O.K. Esteban, and C.E. Borges. *Environment Modelling for Spatial Load Forecasting*. Lecture Notes in Computer Science. Springer, to be published in 2015.

[6] J. Zhao and W. Deng. Relationship of walk access distance to rapid rail transit stations with personal characteristics and station context. *Journal of Urban Planning and Development*, 139(4):311–321, 2013.