Accessibility on a micro-level: a closer look at pedestrian routing and network generation

Michael A.B. VAN EGGERMOND a and Alex ERATH b

a,b Future Cities Laboratory (FCL), Singapore ETH Centre, 138602, Singapore
a E-mail: erath@ivt.baug.ethz.ch
b E-mail: eggermond@ivt.baug.ethz.ch

Abstract: The main purpose of this paper is to investigate ways to determine realistic pedestrian distances from an origin to a destination and to calculate geographic extents based on pedestrian distances. Pedestrian movement and routing research employs a range of methods, requiring different types of data. Pedestrian movement is employed on the microscopic level. Research towards pedestrian networks has focused on generating sidewalks from different types of data sources in large areas. In pedestrian movement and network generation research generalized costs for pedestrians have not been investigated for different user groups. We take methods from movement research and investigate the possibility to employ this on a larger scale for routing purposes. Network generation methods and data sources are compared with each other and ground-truth. For future research we envisage a hybrid approach combining a grid for certain areas with a backbone of sidewalks and different link costs for different user groups.

Keywords: Pedestrian Networks; Pedestrian Movement; Accessibility; Routing

1 INTRODUCTION

Each day individuals are faced with a wide ranges of choices, such as mode choice, destination choice and departure time choice. Dependent on the study region, sample and control variables, the outcome of these choices are dependant on the availability of a car, the availability of public transport, trip purpose, accessibility and the built environment. Especially the influence of the built environment on the outcome of choices has received attention in research in recent years. Cervero & Kockelman (1997) introduced the 3D’s to quantify the relation between travel and built environment: diversity, density and design. More recently, several factors have been added to quantify the relationship between travel and built environment: destination accessibility, distance to transit, demand management and demographics (Ewing & Cervero, 2010). In an extensive meta-analysis Ewing & Cervero (2010) show that mode share and likelihood of walk trips are most strongly associated with the design and diversity elements of built environments. Intersection density, job-housing balance and distance to stores have the greatest impact.

With an ever-increasing part of the population in Asia residing in cities local governments have formulated policies to house this population in high density environments. Examples can be found in the new towns in China, Hong Kong and Singapore. To curb particle emissions and congestion, it becomes important to consider policies to curb private transport and promote other modes, such as walking and public transport. Public transport trips will always constitute of walking at the access and egress stage of such a trip. Quantifying factors such as diversity, density and design is highly relevant for urban planners, transport planners, and policy-makers.

Walking distances between an origin and one or more points or areas of interest can be calculated by either measuring the euclidean distance (multiplied with a detour factor) or by the network distance. Measures such as job-housing balance, population density and shopping opportunities all require a geographic area in which this indicator is calculated. Three main
approaches can be recognized to calculate the geographic area: (1) spatial buffers (circular, rectangular, hexagon, etc.), (2) geographical boundaries (neighbourhoods, traffic zones, planning zones, census districts etc.) and (3) network distances (extent of the area that can be reached in a pre-defined network distance, time or generalized cost).

That different geographic extents are of relevance is for instance shown by Guo & Bhat (2007). They estimate residential location models and shows that people perceive different variables on different levels; i.e. population density is perceived on a more local level than land-use variables. A study by Eng (1994) for Singapore shows that neighborhood perception and awareness, measured in job, shopping, sports and leisure opportunities, depends on canals and main roads, and also by income levels. People with a higher income are more aware of job opportunities located further away. A different example of the relevance of pedestrian networks can be found in the field of retailing: high footfall is preferred by retailers in certain segments (e.g. Löchl, 2008). Commercial examples of the calculation of local accessibility measures to inform potential movers, either being it retailer or future residents, can be found with Walk Score (2012) and Mapnificent (2012).

To truly capture the diversity of the city, but also destination accessibility, it is thus important to consider the walking part of each journey, whether the journey consists of walking, car, public transport or a combination of these modes. However, to be able to calculate the pedestrian distances and costs a pedestrian network is needed. Such a network differs from traditional routing networks for private transport. Pedestrians have for instance different possibilities to cross a road than cars, a sidewalk might not always be available but also totally different networks may exist, such as paths through malls, parks and residential estates. Furthermore, some user groups might not be able to use all elements of a network; i.e. parents with a pram or wheelchair bound folk cannot climb stairs.

Two main elements in literature regarding walking can be recognized: pedestrian routing & movement and pedestrian network generation. Section 2 will continue with different routing principles for pedestrians. From this sections a set of requirements will follow which a pedestrian network should fulfill. Section 3 will discuss network elements and both open source and proprietary data sources for the case of Singapore. It formulates a number of practical considerations a modeler should make when generating a pedestrian network. These considerations can aid future research. Both sections will highlight previous efforts found in literature. Section 4 will finish with future research.

2 PEDESTRIAN MOVEMENT

Vehicle routing in transport is commonly been done with shortest-path algorithms such as Dijkstra. With such an algorithm an element to be routed (car, agent, vehicle) is given a pre-defined origin and destination and the shortest path, expressed in generalized costs, is found. Vehicle behavior has for example been modelled with car-following and lane-changing models. Pedestrians however, unlike cars, are free to move around in different spaces. This has led to different methods to describe the movement of pedestrians in space. Similar to vehicle movement, a differentiation can be made between two broad group that describe pedestrian movement: (1) models considering a fixed origin-destination pair (e.g. Helbing & Molnár, 1995) and (2) models where the pedestrian is given a direction or activity schedule, but considered exploratory rather than purposive (e.g. Penn & Turner, 2001).
2.1 Fixed origin-destination pairs

Models considering fixed origin-destination pairs mostly operate on a microscopic or mesoscopic level. Helbing & Molnár (1995) introduce the social-fore model and apply a physics approach to model pedestrian movement. Different behavioral forces, such as resistance and attraction, describe pedestrian patterns. In this model it is not important which pedestrian performs an action, as fractions of people are considered. This model is for instance applied for pedestrian evacuation dynamics (Helbing et al., 2002). Y.F. Tang & S.C. Pun-Cheng (2004) present a path computational model based on topographic map features. In their model they aim to compute a path not based on the road centre line, but rather use a combination of different data sources, such as street blocks, buildings blocks and intersections. A pedestrian can walk over some of these areas, but is blocked from others. Such an approach comes closer to network routing, as between origin and destination a line is drawn.

2.2 Exploratory models

Instead of focusing on a single origin-destination pair, exploratory models aim to model the movement of pedestrian through space. In this review, some approaches will be highlighted. A first approach has its origins in shopping behavior. A pedestrian has a set of activities to be carried out within a confined space, such as a shopping mall, street or urban centre and the execution of this schedule is observed. These models contain information about the familiarity with the shopping area as well as socio-economic characteristics of the pedestrian. Data for such models stems from both questionnaires and individual tracking devices. An example is the STREETS pedestrian model (Schelhorn et al., 1999) pedestrians movement is determined by spatial configuration, local attractions and activity schedules. Borgers et al. (2009) provide an extensive literature review concerning models pedestrian movement in a shopping environment. A second approach is presented by Kukla et al. (2002). In their model, the PEDFLOW model, pedestrians are modeled autonomous agents moving through a virtual environment. The virtual environment can be seen as a set of grid cells, which either are available or blocked. Pedestrians only have limited local knowledge of the system. Penn & Turner (2001) apply a space syntax approach to pedestrian movement and give agents a set of pre-processed information of the system. Agents employ a visibility graph to select a next destination from a set of possible locations.

2.3 Required elements for a network

The main interest of this paper is to investigate possible ways to determine realistic pedestrian distances (or costs) from a certain origin to a destination or to calculate geographic extents based on pedestrian distances. Which lessons can then be drawn from previous research? First, we see that interaction with other pedestrians has received attention to capture crowd and evacuation dynamics. Second, Y.F. Tang & S.C. Pun-Cheng (2004) present a model which matches built environment and network elements, which offers a promising approach for calculating shortest paths. Third, attention has been paid to how pedestrians traverse a space, based on a combination of their socio-economic characteristics and activity-schedule. Not much attention has been paid to other elements that might influence pedestrian behavior and routes have been considered in this brief literature review, such as stairs or traffic levels on a road. Instead, it is assumed an
element of the network is either accessible or blocked. Space syntax visibility graphs might indicate how attractive a certain area is to walk by; as well as pedestrian movement in a shopping settings. Factors that might influence pedestrian movement, such as traffic levels, pavement type, shelter from elements, slopes (stairs or ramps) have not been analyzed. These factors are of interest, as they can create different pedestrian routes for different socio-economic groups. Examples are abound: a wheelchair bound person is not able to use a pedestrian bridge but requires an equal level intersection; tourists posses different network knowledge than locals. In the next section we will discuss pedestrian network generation from different data sources.

In this section it could be seen that a pedestrian network is not always required. Dependent on the routing or path finding algorithm chosen different elements of the built environment are required. In the path-finding algorithm presented by Y.F. Tang & S.C. Pun-Cheng (2004) a distinction is made between street blocks (blocks between streets, building blocks, links (bridges, zebra-crossings) and roads. Kim et al. (2009), focusing only pedestrian network generation, list a number of elements, such as streets, sidewalks, overhead bridges, pedestrian crossings, underpasses and building entrances.

Table 1 lists the aforementioned elements and several others. The spatial representation of these elements in spatial datasets is listed in the first column. The remaining columns indicate the spatial representation of the elements for two types of routing: network based routing and exploratory movement. In network based routing much information of the surrounding is lost, as buildings and parcels are not available anymore but are simplified to points linked to the network through entries and exits. On other hand, the representation is exploratory movement is largely based on polygons, which might need to be created from different types of input data. Crossings, mostly represented as lines in spatial data sets, will need to be exploded to polygons. From this table it is not clear whether elements such as shelter, shade, traffic levels, steepness and pavement width are encapsulated as attributes of the different elements. All these attributes contribute to the generalized costs of a pedestrian trip.

<table>
<thead>
<tr>
<th>Element</th>
<th>Representation in spatial datasets</th>
<th>Network routing</th>
<th>Exploratory pedestrian movement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roads</td>
<td>Line</td>
<td>Line</td>
<td>Polygon</td>
</tr>
<tr>
<td>Sidewalks</td>
<td>Line</td>
<td>Line</td>
<td>Polygon</td>
</tr>
<tr>
<td>Land parcels</td>
<td>Polygon</td>
<td>-</td>
<td>Polygon</td>
</tr>
<tr>
<td>Parcel entry and exits</td>
<td>Point / Line</td>
<td>Point</td>
<td>Polygon</td>
</tr>
<tr>
<td>Buildings</td>
<td>Polygon</td>
<td>-</td>
<td>Polygon</td>
</tr>
<tr>
<td>Buildings entries and exits</td>
<td>Point / Line</td>
<td>Point</td>
<td>Polygon</td>
</tr>
<tr>
<td>Building footpaths</td>
<td>Line</td>
<td>Line</td>
<td>Polygon</td>
</tr>
<tr>
<td>Overhead bridges</td>
<td>Polygon</td>
<td>Line</td>
<td>Polygon</td>
</tr>
<tr>
<td>Underpasses</td>
<td>Polygon</td>
<td>Line</td>
<td>Polygon</td>
</tr>
<tr>
<td>Car intersections</td>
<td>Points</td>
<td>Point</td>
<td>Polygon</td>
</tr>
<tr>
<td>Crossings</td>
<td>Polygon</td>
<td>Line</td>
<td>Polygon</td>
</tr>
<tr>
<td>Traffic lights</td>
<td>Points</td>
<td>Point</td>
<td>-</td>
</tr>
<tr>
<td>Parks</td>
<td>Polygon</td>
<td>-</td>
<td>Polygon</td>
</tr>
<tr>
<td>Rivers/ lakes</td>
<td>Polygon</td>
<td>-</td>
<td>Polygon</td>
</tr>
<tr>
<td>Points of interest</td>
<td>Line</td>
<td>Point</td>
<td>Point</td>
</tr>
<tr>
<td>Pedestrian zones (squares)</td>
<td>Polygon</td>
<td>Line</td>
<td>Polygon</td>
</tr>
</tbody>
</table>
3 PEDESTRIAN NETWORK GENERATION

Whereas pedestrian movement has received attention from numerous disciplines, research towards pedestrian network generation has been conducted mainly within the field of Geographical Information Sciences (GIS). Karimi & Kasemsuppakorn (2012) present an extensive review of pedestrian network map generation approaches. A distinction can be made between three main approaches: network buffering, collaborative mapping with GPS traces and image processing. Of these approaches, network buffering requires the least computation effort; image processing most. The outcome of all three approaches depends on the completeness and quality of the input data. Ballester et al. (2011) present a method based on images to generate sidewalks and walkable plains (i.e. squares) and see as possible improvements the inclusion of park paths, bridges and pedestrian crossings. Kim et al. (2009) present a method based on existing spatial sets and make a distinction between roads and crossings and point out that the accuracy of the pedestrian network is strongly based on the sophistication of the spatial datasets. However, whether they create sidewalks from road centre lines or have sidewalks available digitally is not clear. Parker & Vanderslice (2011) propose a tool to create sidewalks from road centre lines and create crossings based on road intersections. In the proposed tool it is possible to exclude certain road categories from having sidewalks (e.g. expressways).

3.1 Data sources

The elements required for a pedestrian network usually are stored in data sets which may or may not come from the same data suppliers. The advantage of open source data, such as OpenStreetMap (OpenStreetMap, 2011) is that it is available freely for researchers. The downside is that accuracy is not guaranteed. Also, accuracy might not be the same spatially; some areas might contain more detail than other. Transport, planning and land authorities maintain their own data sets. Finally, private companies, such as navigation map providers will have their own maps. In Figure 1 to Figure 3 maps from these three providers are depicted for a randomly chosen area in Singapore. As reference Google Maps is chosen. Figure 1 contains data from the Singapore Land Authority (SLA). It can be seen that the geometry of the roads as on Google is similar. Figure 2 contains data from OpenStreetMap. It can be seen that this figure contains more lines; this can be due to the format in which OpenStreetMap is stored. All lines for Singapore are contained in a single file, from which different categories can be filtered, such as bridge, steps, footway and highway. Closer inspection of SLA data has shown us that it also contains non-navigable roads by cars, for instance in parks. OpenStreetMap data for Singapore is not complete when selecting the category footways. Furthermore, in this example some sidewalks are not captured, as only one road centre line is drawn (Figure 2, left side). Both SLA and OpenStreetMap contain overhead bridges. Figure 3 contains roads from a NAVTEQ navigation network (NAVTEQ, 2011). Here it can be seen that the geometry differs from the geometries. One could argue that this need not to be a problem if distance attributes were similar. However, if in a next step an individual would walk from a building to the NAVTEQ network these distances would not represent real-life distances.

One example of the pedestrian network captured in OpenStreetMap can be seen in Figure 4. Steps leading from a residential estate to a bus stop are drawn. Furthermore, the attribute name 'steps' would make it possible to exclude certain user groups from this pedestrian way.
3.2 Pedestrian network generation for Singapore: initial steps

To test some of the concepts described in the previous sections we take Singapore as a case-study. This selection is one hand based on the availability of different types of data sources for the entire island as well as familiarity with the real-life pedestrian network.

Network buffering

As a first step network buffers have been generated for all roads in SLA road network. Crossings have been removed from the resulting network. In this way a possible set of sidewalks is generated. In this initial step expressways are initially not accounted for; they can still have sidewalks with this crude approach. However, as sidewalks still have their original ID and the SLA network contains a road classification, this should be easy to resolve in a later stage. In first instance, we do not assume pedestrians are able to cross at all road intersections. To these means, we use a data set containing the location of traffic lights, including pedestrian lights. This data set is not connected in a relational way to the road network. Therefore, a rule-based approach is required to create pedestrian crossings at these intersections. Pedestrian overhead bridges serve as a connector between sidewalks separated by a road centre line. Less easy to resolve is how many lanes would need to be crossed by a pedestrian if no overhead bridge or traffic lights are available. The number of lanes could be an indicator for (1) negative perception for crossing a road at a certain point and (2) the possibility to jaywalk. At this moment the
possibilities are still being investigated. A second case where crossing is not possible is where railings have been installed next to the sidewalk or in the road centre. Figure 5 depicts the generated sidewalks, the modeled pedestrian overhead bridge and ground truth. This map section is chosen as it depicts one downside of using road centre lines: road centre lines, not being an expressway, might not have sidewalks. In this case the middle part of the road enters a tunnel whereas local traffic can turn left. The overhead bridge, as currently modeled, connects all generated sidewalks, also the ones not present in real-life. One could argue that this doesn’t influence pedestrian routing, as pedestrians will eventually reach a dead-end. In a shortest-path this route would thus not be considered.

One important point to consider with the network buffering approach are building entries/exits and parcel entries/exits. This is illustrated in Figure 6. The green route indicates a building exit and parcel exit to the right side, leading to a longer distance to the bus stop. The red line indicates a possible rear exit of the parcel, leading to a much shorter distance. A similar
problem can hold for matching points of interest to the network; points of interest, such as bus stops and subway entries need to be assigned one or more points on a link.

*Grid-cell approach*

Notwithstanding the fact that the grid-cell approach as proposed by Kukla et al. (2002) was meant for exploratory movement by pedestrians, the grid-cells do offer the opportunity to define a walkable plain. Furthermore, it is possible to define a different plain for different user groups by giving each grid cell different costs. As example we have considered the area in Figure 1, which measures approximately 900m x 900m. Each gridcell measures 1m x 1m. The walkable plain is defined as space with no built up area (roads, parcels with buildings). Thus, we are limiting the definition of walkable plain as given by Penn & Turner (2001) as we also limit access to parcels and not only to buildings. Figure 7 shows the result of the grid-cell approach. Walkable is considered to be the yellow area; areas coloured green can not be traversed by pedestrians. Note that pedestrians are able to traverse the steps depicted in Figure 4. However, a building underpass For the case of Singapore, special attention needs to be paid to parcels containing public housing. This area is publicly accessible, as opposed to parcels containing commercial buildings and private housing. As this data is available to the researchers, this is envisioned in a next step. With the grid-cell approach it is also necessary to define entries and exits to parcels. Furthermore it is necessary to define ‘inside’ and ‘outside’ a parcel. Pedestrians ‘inside’ a parcel are possible to leave a parcel; pedestrians ’outside’ a parcel are not able to enter except if the parcel is their destination.

One downside of this approach is the computational effort. In order to compute shortest paths over such a grid is both time and memory intensive. The required data to describe the network in this small part of Singapore is many times larger than the entire navigation network of Singapore.

*Figure 7: Walkable plain on a 1m x 1m grid; yellow being accessible by pedestrians, green and grey are not accessible (left). Ground-truth building underpass, location indicated by the arrow in the left image (Google, 2010)*

We envisage a second hybrid approach - the grid-network approach. Parcels will be categorized containing commercial buildings and private housing. This area is publicly accessible, as opposed to parcels containing commercial buildings and private housing. As this data is available to the researchers, this is envisioned in a next step. With the grid-cell approach it is also necessary to define entries and exits to parcels. Furthermore it is necessary to define ‘inside’ and ‘outside’ a parcel. Pedestrians ‘inside’ a parcel are possible to leave a parcel; pedestrians ’outside’ a parcel are not able to enter except if the parcel is their destination.

*A hybrid approach - the grid-network approach*

The path-arc model proposed by Y.F. Tang & S.C. Pun-Cheng (2004) can be considered a hybrid approach. In their approach, they construct a line between source and destination and evaluate intersection points with polygons on the line. From the first intersection, again a set of lines is drawn. This approach will investigated further in the near future.

We envisage a second hybrid approach - the grid-network approach. Parcels will be categorized containing commercial buildings and private housing. This area is publicly accessible, as opposed to parcels containing commercial buildings and private housing. As this data is available to the researchers, this is envisioned in a next step. With the grid-cell approach it is also necessary to define entries and exits to parcels. Furthermore it is necessary to define ‘inside’ and ‘outside’ a parcel. Pedestrians ‘inside’ a parcel are possible to leave a parcel; pedestrians ’outside’ a parcel are not able to enter except if the parcel is their destination.

*One downside of this approach is the computational effort. In order to compute shortest paths over such a grid is both time and memory intensive. The required data to describe the network in this small part of Singapore is many times larger than the entire navigation network of Singapore.*
according to their permeability to the pedestrian: parks can be traversed freely, shopping malls have a boundaries have a lower permeability as it is not clear where the entrances might be, public residential areas are accessible and private property is not accessible to pedestrians. These rules will be defined for different land-use types and building types. From permeable parcels a grid will created. Link costs of this grid will be based on the permeability and will be different for different user groups. As backbone for this hybrid grid-network approach a set of generated sidewalks will be taken. Grids will be connected to this network over their respective interior grids. Such an interior grid might take the form of a directed graph, to make leaving possible, or of an hierarchical network. In this way we aim to ensure accuracy but also aim to make the network more efficient.

4 OUTLOOK

The main interest of this paper is to investigate possible ways to determine realistic pedestrian distances (or costs) from a certain origin to a destination or to calculate geographic extents based on pedestrian distances. Pedestrian movement and routing research employs a wide range of methods, requiring different types of input data. Mostly, these methods are employed on the microscopic or mesoscopic level. Research towards pedestrian network generation has mainly focused on generating sidewalks and crossings from different types of data sources in large areas, such as cities and countries.

In the near future we will implement the grid-network approach and evaluate it’s performance as compared to a more advanced path-arc model and Euclidean distances. Also, as open-source map data contained more elements as expected, we will investigate this strand of research further. For both the hybrid grid-network approach and a purely network based approach it remains necessary to model building entries and exits and parcel entries and exits. This will also be investigated further. In both approaches a full generalized cost approach will be considered. As such, a more elaborate list of network elements and their attributes will be given in future work.

A second application for a pedestrian network can be found in the state-of-the-art agent- and activity-based transport demand models. Agent-based MATSim (2012) is one example of a model that puts the activity-based approach in practice and contains a fully integrated traffic flow simulation to calculate the generalized costs implied by the activity schedule. In such a model, each individual in the population is considered an agent; the number of activity locations depends on the level of granularity considered. The implementation of MATSim for Zurich considers parcels of 500m by 500m; the implementation of MATSim for Singapore considers the individual building (Erath et al., 2012). Such a model yields temporal differentiated travel times for both public transport and private transport from all buildings to all buildings and can aid in calculating accessibility on a micro-level. If daily choices, such as mode choice and destination choice are to be modeled on the level of the individual building, a form of pedestrian network is required.

However, what we consider most important is that access and egress to different modes and different user groups should be accurately incorporated in accessibility calculations on a micro-level. This can aid transport planners, architects and policy-makers in designs, plans and policies in which the pedestrian is accounted for.
ACKNOWLEDGEMENTS

The research conducted at the Future Cities Laboratory is funded by the Singaporean National Research Fund. We wish to express our gratitude to the Land Transport Authority for providing us invaluable data sets on transport in Singapore. Also we are very thankful to the Singapore Land Authority for providing us with a wide range of data sets.

REFERENCES


