On the Road to Nowhere? Least Cost Paths, Accessibility and the Predictive Modelling Perspective

Verhagen, P.
CLUE, Faculty of Arts, VU University, Amsterdam, the Netherlands
jwhp.verhagen@let.vu.nl

This paper reports the first results of a study into the utility of using least cost path models and other indicators of accessibility for predictive modelling purposes. While accessibility and movement potential are potentially important variables influencing settlement patterns, a number of difficulties are associated with the application of cost surfaces, least cost paths and network analysis for this purpose. The paper discusses the relevant issues, and presents a new way of creating maps of regional movement potential through the use of cumulative cost paths. It is concluded that a better theoretical foundation of (pre-)historic movement is necessary to apply these techniques more successfully for the analysis and prediction of settlement patterns.

Keywords: least cost path modelling, accessibility, predictive modelling.

1. Least cost paths, accessibility and predictive modelling

In archaeological predictive modelling, it has always been considered difficult to include socio-cultural factors in the models used (VERHAGEN et al., 2010). One of these factors is the movement of people, animals and material resources through the landscape. Every archaeologist is aware that transport routes may have had a major influence on the density of settlement and the accumulation of archaeological materials; and the routes themselves are archaeological sites as well. From a predictive modelling point of view, least cost path (LCP) modelling may therefore be an important tool that can contribute to the prediction of locations or zones where (pre-)historic movement patterns may have concentrated. Yet, LCP modelling and territorial modelling based on cost surfaces have not been used extensively for predictive modelling purposes (but see WHITLEY and BURNS, 2008; WHITLEY et al., 2009). Most published studies consider the calculation of possible paths between known archaeological sites (e.g. VAN LEUSEN, 2002; HOWEY, 2007; ZAKŠEK et al., 2008), or even try to reconstruct (partly) known routes by means of LCP modelling (e.g. BELL and LOCK, 2000; FIZ and ORENGO, 2008; POLLA, 2009). The cumulative cost surfaces that form the basis for LCP modelling are used regularly to model settlement territories (e.g. SOETENS et al., 2003; ROBB and VAN HOVE, 2003; DUCKE and KROEFGES, 2008).

The debate in archaeological computing literature has mainly centred on questions concerning the correct ways of defining the friction surfaces used, especially where it comes to the impact of slope on movement, and on the algorithms used for calculating the LCPs (e.g. LLOBERA, 2000; EJSTRUD, 2005; ZAKŠEK et al., 2008; GIETL et al., 2008; HERZOG and POSLUSCHNY in press; HERZOG, 2010).

LCP modelling however is only one of the techniques available to model and measure connections between geographic locations. In economic geography, network analysis is used extensively for the same purposes (see e.g. RODRIGUE et al., 2009). Space syntax is a rapidly growing field of research dealing with similar issues. It is usually applied more in urban and built-up contexts (see BAFNA, 2003 for an introduction, and e.g. CRAANE, 2009 for an historical case study). Both approaches provide powerful tools to extract measures of connectivity between network nodes (places of departure and arrival). Measures of ‘resistance’ (friction) along the edges of the graph (the paths) can be used as well to find the travel routes with the least costs, especially over longer distances. The number of landscape archaeological applications of network analysis is however rather limited (see e.g. ALDEN, 1979; HARE, 2004; NUNINGER et al., 2006), and it has never been considered for predictive modelling purposes. Cost surfaces, LCP modelling and network analysis can be useful tools to identify places that are inherently

Issues in Least-Cost Analysis
more connected or isolated than others, which might tell us something about the attractiveness of certain portions of the landscape for settlement and other activities. This can also be analysed at different scales (see LLOBERA, 2000). For example, a hill-fort is usually located in a position that is difficult to reach from a short distance, yet it can at the same time be in or close to an area attractive to movement.

Surprisingly, a methodological connection between the raster-based LCP approaches and vector-based network analysis tools is lacking, even within disciplines like transport economics. LLOBERA (2000) approached the issue of landscape accessibility with raster GIS and used the concept of total path costs to obtain indices of accessibility of the landscape at various scales. His approach is based on a network analysis technique known as the construction of a Shimbel Distance or D-matrix, or all-pair shortest path matrix that holds the shortest paths between all nodes of a network (SHIMBEL, 1953). It gives the average cost of moving from one location to all other locations in the network. If all nodes are at equal distance, then the best connected places will be those in the centre of the region considered. However, as soon as distances between nodes start to differ (in terms of real distances and/or in terms of friction) a pattern of differential accessibility of nodes will emerge, depending on the maximum movement distance chosen. Such a matrix can for example be calculated with the space syntax software package Depthmap (http://www.vr.ucl.ac.uk/depthmap), where it is called the total depth - defined as the cumulative total of the shortest paths between the line segments of a network (TURNER, 2004).

From a network analysis perspective raster maps are just collections of evenly spaced nodes. The construction of a Shimbel Distance matrix for a raster will therefore only create zones of differential accessibility when introducing friction. The configuration of these zones will then depend on the maximum movement distance chosen. Unfortunately, the construction of Shimbel Distance matrices for raster maps is time-consuming since it is necessary to calculate and cumulate all possible connections for each individual raster cell. LLOBERA (2000) therefore developed a custom-made routine for this that at the time was not integrated in a GIS platform. MLEKUŽ and VERMEULEN (2010) show a method for calculating this kind of accessibility maps that can be applied with relative ease in raster GIS for different movement distances. However, this still takes a considerable amount of time to calculate, and in practice a quicker solution is to apply a neighbourhood mean or sum filter on the friction surface used, using the desired movement distance to determine the size of a circular neighbourhood radius. While the actual numbers obtained this way are not directly comparable to a Shimbel Distance matrix, this method will provide comparable images of zones of differential accessibility.

2. From accessibility to movement: creating cumulative cost paths

The calculation of accessibility maps from Shimbel Distance matrices, or by using simpler methods, does not give us direct information on potential movement patterns. A location might be relatively inaccessible, yet it could make a large difference for movement patterns whether it is difficult to reach from all sides, or just from a limited number of directions. The modelling of potential movement corridors is an issue that has received little attention in archaeology. The main application is found in ecology, where corridor models are routinely used to predict animal movement. An interesting approach in this respect is the use of principles from electronic circuit theory to predict movement patterns (MCRAE et al., 2008) with the open source Circuitscape software (http://www.circuitscape.org). However, these ecological corridor models are always used in a way that is also common in archaeology, only using known points of departure and arrival. This is not necessarily what we are interested in from a predictive modelling perspective, since in most cases we don’t know the actual distribution of settlements and other points of interest where people may have moved from or to.

By combining the ideas discussed above, it is nevertheless possible to create maps of regional movement potential. An example of this is already found in the case study by WHITLEY and BURNS (2008) who modelled potential movement routes of hunter-gatherers through a study area in South Carolina, simply by calculating LCPs starting at 1 km intervals at one edge of the study area, and ending at the other edge. This approach does not take any preconceptions of settlement structure into account. The landscape is seen as containing a number of potential pathways, some of which are more likely to be used than others, depending on the friction surfaces defined.

We can take this approach one step further by calculating multiple LCPs for a reasonably large number of “non-site” sample points in a region. Multiple cost paths to each sample point within a predefined radius of movement can then be calculated and added together, like ZAKŠEK et al. (2008) did for a set of cost paths obtained for known settlement locations. This will then result in a map of LCP densities for the whole region that is considered.

3. Application

This idea was tested in a study area in the east of the Netherlands, the Rijssen-Wierden region, measuring 10 by 12 km. In this particular region, elevation ranges between 7 and 37 m above sea level, and, while some low hills are found, slope cannot be considered a major obstruction to movement. Water however certainly is. Substantial parts of the area were covered by peat marshes until the beginning of the 20th century, and
movement through these areas would have been slow and treacherous.

As the case study was only undertaken for purposes of experimentation, no attempt was made to reconstruct the spatial distribution of marshes and fens in great detail. Instead, the groundwater table registered on the 1:50.000 soil map of the area was used as a rough approximation of the wetness of various portions of the landscape. Using this information, three basic categories of accessibility were defined: dry areas, humid areas and wet areas, with corresponding friction factors of 1, 2 and 4 (figure 1). Basically, this means that crossing a wet area would take 4 times as long as crossing a dry piece of land.

In order to model preferred pathways, 120 points spaced 1 km apart were placed in the study region. For each point a cumulative cost surface of moving to this location was calculated, using the \texttt{r.cost} module in GRASS 6.4. Then, paths moving to these locations were calculated from 72 starting points radially distributed around the sample points, using the \texttt{r.drain} module in GRASS. The paths calculated for each location were then added together into what we could call a cumulative cost path (CCP) map. This was done for ‘travel distances’ of 250, 1000, 2500 and 5000 meters. At 5000 meters the distance covered is almost the complete study region, so no larger radii were used. The paths were allowed to run out of the study area if the edge was closer than the radius distance. In practice, this means that paths close to the edge of the region look like paths that are created using shorter travel distances.

4. Results

Visual inspection of the experiments’ results shows that the method creates paths that cross the entire area. When using a 5000 m travel distance these paths tend to converge at places where wet areas can be crossed with relative ease (figure 2). For example, clear crossings are seen where valleys are relatively narrow, and paths tend to prefer to follow the edges of the dry zones to reach those crossings. Furthermore, a network of preferred routes is clearly visible in the east part of the area, where small pockets of dry land (sandy ridges) are connected. In larger areas of dry land, the paths tend to become less clear, and follow direct, straight routes, partly as a consequence of using queen’s move instead of knight’s move for creating the cost surfaces. At smaller travel distances, the networks become much less clear (figure 3). These results conform to logic: for long distance travel the effect of avoiding areas of low accessibility is larger than for short distances, and deviations from ‘random’ paths will then accumulate in places with relatively good accessibility, creating corridors of preferred movement.

The CCP approach takes the number of times a LCP is created in each of the simulations as the basic indicator of accessibility, and in that way creates a network that exhibits differences in path density depending on the sample locations used and the distance of movement chosen. Adding or removing starting or end points, or using random instead of evenly spaced points did not seem to fundamentally change the structure of the network created. The highest path density is always found in places where travel is forced through corridors of relatively easy movement within areas that are more difficult to negotiate. In places where movement is not strongly reduced, paths may go in any direction.

The maps created can, with some further manipulation, be integrated with other network analysis techniques like space syntax. Figure 4 shows an example of this. In order to obtain a vector network from the modelled CCPs, one additional step was taken. A random noise factor was introduced to the friction surface to obtain CCPs that show a stronger concentration of paths in homogeneous areas. The CCP-map was then vectorized and imported in Depthmap, where the choice parameter (the number of times a location is encountered on a path from origin to destination) was calculated as an indicator of the potential for through-movement (TURNER, 2004).
5. Discussion

There are a number of issues that still need to be resolved before we can use LCP modelling and network analysis to better effect for predictive modelling purposes. First of all, no generally accepted methods are available to analyse and quantify the differences in outcome between the various LCP models. One of the few methods found in literature is the detour index (RODRIGUE et al., 2009). This only specifies the relative deviation of one route from a different one, and not the differences between various cost surfaces or cumulative cost paths. EJSTRUD (2005) used quadrant analysis to calculate the level of agreement between different cost path models; this is a more versatile method, as it can take multiple paths into consideration.

Secondly, computing time still is an issue for calculating the cumulative cost paths presented here. These were calculated using a limited amount of starting and end points. This seems to be justified, as increasing the number of points used did not drastically alter the patterns observed. However, a ‘total cost path’ map, with cost paths modelled for each raster cell in the area would still take a considerable amount of time to calculate.

Apart from that, we need to have a closer look at how to interpret the results of the CCP modelling. The evidence for prehistoric travel is on the whole relatively scarce.
since it has not left many permanent traces in the landscape, except in cases where we have remnants of paved roads and bridge construction. Consequently, proving or disproving the existence of trails and paths will be impossible in most cases, and only theoretical considerations concerning the creation and maintenance of paths can be applied. At the very best we can hope that the modelled paths are close enough to reality to allow us to better locate attractive zones for settlement or other activities.

Unfortunately, (anthropological) theory on how and why (pre-)historic paths were chosen and maintained is not very well developed and LCP modelling seems to be almost ignored by archaeologists interested in roads and trails, as is witnessed by most of the papers in SNEAD et al. (2009). BECKER and ALTSCHUL (2008) however tested ethnographic travel models with archaeological data and were able to use the results for predictive modelling purposes. At the other end of the scale, the methods from theoretical biophysics proposed by LLOBERA and SLUCKIN (2007) to model up- and downhill movement are very complex to implement in GIS and still lack application in an archaeological context. Most researchers will therefore settle for relatively simple friction specifications based on expert judgement, like the ones applied by HOWEY (2007). Consequently, translating the results of CCP models (or other approaches to accessibility) into a predictive map is not straightforward. As mentioned, proximity to a zone with a high potential for movement could be an important settlement location factor. The utility of this assumption however strongly depends on the theoretical foundation of the model used. A basic approach would be to depart from a friction surface based on travel time, and analyse the resulting path densities within different search radii. We have to realize however that we are, even in this simple case, then dealing with ‘stacked’ multi-scalar analyses: the results of the CCP model depend on the travel distance chosen, but the analysis of proximity to the most attractive zones for movement also includes distance. This can pose difficulties for the interpretation of statistics and patterns, and requires a careful analysis of the models for a number of scales.

This is only complicated if we try to incorporate other factors like visibility (ZAKŠEK et al., 2008), weather conditions, different means of transport, or the proximity to resources and settlements. Figure 5 for example shows an attempt to include the known archaeological site pattern into the CCP model. The friction surface is modified by taking into account the travel cost to the known archaeological sites, and multiplying this with the original cost surface. This method creates clear routes between settlements that are at the same time conforming to the restrictions offered by the landscape. Furthermore, we can see ‘roads to nowhere’, that extend from the settlement to areas where no archaeological sites are known. The patterns created are clearly different from those obtained for the ‘siteless’ CCP maps, but some similarities can be observed as well. Intuitively, this map seems very attractive as a representation of possible prehistoric travel routes. From a predictive modelling point of view however, it will only be valuable if we can establish that the patterns created with a subset of the archaeological data have predictive power for the rest. Within the context of the current study, such an approach still needs to be pursued.

Conclusions

The CCP modelling approach presented here constitutes a pragmatic and intuitive way of obtaining maps of potential movement patterns in a landscape, with or without the inclusion of known settlements. It can be executed relatively quickly, using nothing more than a few lines of scripting in GRASS. The models can be run at different scales of movement and result in interpretable images of potentially preferred pathways. The models are however not based on sound theories of (pre-)historic movement patterns and are subject to all other problems associated with least cost path calculations. More importantly, there are no methods available yet for testing and comparing different model scenarios. The approach can therefore not be applied to real-world predictive modelling questions yet. It is hoped however that this paper provides a first step in this direction.
Acknowledgements

The research for this paper was done within the framework of the research project “Introducing the human (f)actor in predictive modelling for archaeology” (VENI-grant awarded by NWO, The Netherlands Organisation for Scientific Research). The author would like to thank Irmela Herzog, Marcos Llobera and Dimitrij Mlekuž for their valuable and insightful comments that helped to improve the paper.

References


