Relative Neighbourhood Networks for Archaeological Analysis

Diego Jiménez-Badillo
National Institute of Anthropology and History, Mexico. *diego.jimenez61@gmail.com*

Abstract:
This paper describes a method for discovering meaningful relationships in a point set, based on the concept of relative neighbourhood, which can be applied to groups of archaeological sites or artefacts. In contrast to notions that use linear distances, such as nearest neighbour, the relative neighbourhood concept explores "regions of influence" belonging to pairs of points. This allows extracting several types of proximity networks using only the spatial coordinates of the point set. As explained in the paper, relative neighbourhoods can reveal more effectively the contextual relationships of sites or artefacts. The paper also explains how to measure two connectivity properties called integration and control, which are useful to understand the degree of accessibility and relative importance of each node within the network structure.

Key Words: *Archaeological Network Analysis, Archaeological Spatial Analysis, Relative Neighbourhood, Graph Connectivity*

Introduction

In many archaeological projects one is given the spatial coordinates of a set of points (e.g. sites, artifacts) and is asked to discover relational links, such that the resulting network becomes perceptually meaningful in some sense (Fig. 1).

In the case of archaeological sites, for example, we may want to explore several scenarios of inter-settlement relationships to figure out if they reveal some kind of regional organization. Once the network is drawn, we may want to identify sites that exercise a certain degree of control over the network structure, indicating perhaps that their position is more strategic than others. Alternately, we may want to measure how integrated/segregated each settlement is within the network and therefore within the region. After measuring control and integration, we then may use additional information to formulate heuristic hypotheses about the role and importance of each site within past cultural systems. As a way to tackle those issues, this paper:

- Presents a method to derive a relational network purely from the spatial coordinates of a point set.

- Suggests measuring two connectivity properties in order to analyse the degree of accessibility and relative importance of each node of the network in an objective way.

The focus is on the method itself, not on its applications, which have been discussed in other publications (Jiménez-Badillo 2004; 2006; 2009a; 2009b; Jimenez and Chapman 2002).

The approach relies on the graph theoretical notion of relative neighbourhood and therefore belongs to a perspective of spatial analysis based on morphological, topological concepts. This is different from the statistical analysis of point distributions. Given its spatial nature, the approach is also different from analyses focused on virtual links as in social network applications of graph theory (Nooy et al. 2005).

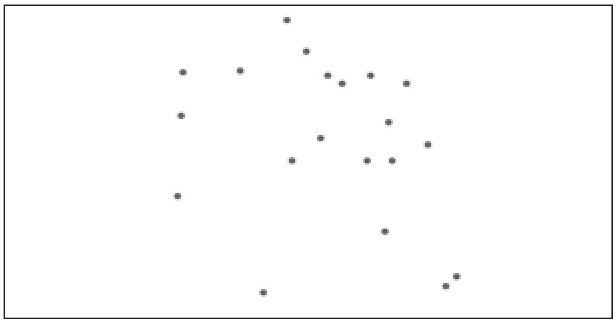


Figure 1. Many projects study point sets to find meaningful relationships between archaeological sites or artifacts. An interesting method is to derive networks purely from their spatial arrangement using the concept of relative neighborhood.

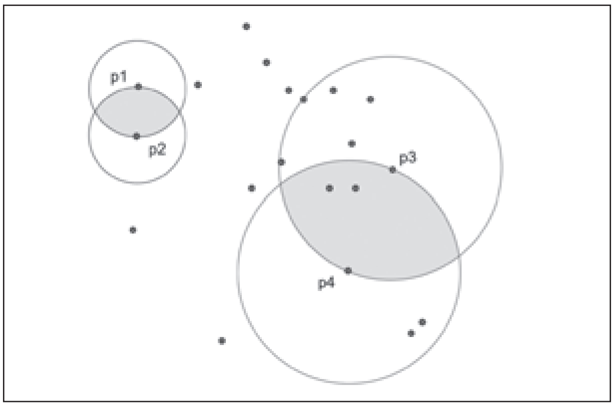


Figure 2. The so-called test of "region emptiness" used to discover meaningful relationships between points. Two sites are said to be relative neighbours if, and only if, their region of influence (shaded area) is empty. In this example, p1 and p2 are relative neighbours while p3 and p4 are not.

The Relative Neighbourhood Concept

As its name suggests, the concept of relative neighbourhood captures the idea of points being "relatively associated" or "more or less related". It has applications in situations where we need to establish contextual relationships of one point, say p_i, with several adjacent points. In other words, when we want to determine whether p_i has other significant spatial associations besides its nearest neighbour. In archaeological jargon we could refer to this as finding different degrees of "association" between sites or artefacts.

It is worthwhile to clarify the difference between the relative notion of neighbourhood and the absolute concept implicit in nearest neighbour. Nearest neighbour considers the location of one object against all the others to determine which item is the closest. To this end it measures linear distances. In contrast, the concept of relative neighbourhood associates "areas of influence" to pairs of points. The size and shape of such regions are determined by the relative separation among all possible pairwise permutations of points and therefore the connections are totally dependent on the morphological configuration of the point set. Approaches like Central Place Theory also involve drawing areas of influence around sites, but consider one region for each single site. In contrast, the relative neighbourhood concept defines regions of influence belonging to pairs of points.

Extraction of Relative Neighbourhood Networks

Imagine a point set $P = \{p_1, p_2, p_3 \dots p_n\}$ and take any two elements, say p_i and p_j. We say that p_i and p_j are *relative neighbours* if, and only if, both are at least as near to each other as they are to the rest of the points. In order to determine whether p_i and p_j comply with such definition it is necessary to perform the following test, known as 'region-emptiness':

- The first step is to take the distance $d(p_i, p_j)$ as the radius for drawing two circles C_i and C_j, centered at p_i and p_j respectively. The intersection of both circles delimits an almond-shaped region called R_{RNG} (*vesica piscis* in Latin), which represents the region of influence for that particular pair. The concept can be applied to point sets located in 2, 3 and higher dimensional space.

- The second step is searching for other members of P within the region R_{RNG}. If R_{RNG}

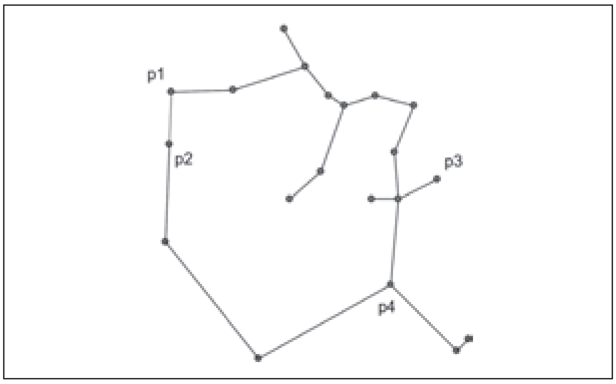


Figure 3. Relative Neighbourhood Graph for the point set illustrated in Figure 2. This is extracted by applying the test of "region emptiness" to every combination of pair of points.

is empty (i.e. no other point of P lies inside), then p_i and p_j are relative neighbours and a line is drawn to join them. On the contrary, if any other element of P, say p_n, lies within R_{RNG}, then p_i and p_j are declared non-neighbours. Figure 2 illustrates both cases. When every possible permutation of points is tested in the above way we obtain the so-called Relative Neighbourhood Graph (RNG). Notice how the RNG matches the visual perception of the point set topology (Figs 3 and 4).

The analysis of "relatively close" points dates back to 1969 when Lankford (1969) defined the concept mathematically. Then, others developed the idea further (Gabriel and Sokal 1969; Matula and Sokal 1980; Kirkpatrick and Radke 1985; Toussaint 1980a, 1980b, 1980c, 1988; Urquhart 1980, 1982, 1983). Besides the Relative Neighbourhood Graph, these efforts produced other proximity graphs such as the Gabriel Graph, the Beta-skeletons, the Limited Neighbourhood Graph, etc. Here, we refer to such graphs collectively as "proximity networks". Simultaneously, several algorithms have been developed to compute these graphs (Huang 1990; Hurtado et al. 2001, Jaromczyk and Toussaint 1992; Jaromczyk and Kowaluk 1987, 1991; Rao 1998; Su and Chang 1990, 1991a, 1991b; Toussaint and Menard 1980).

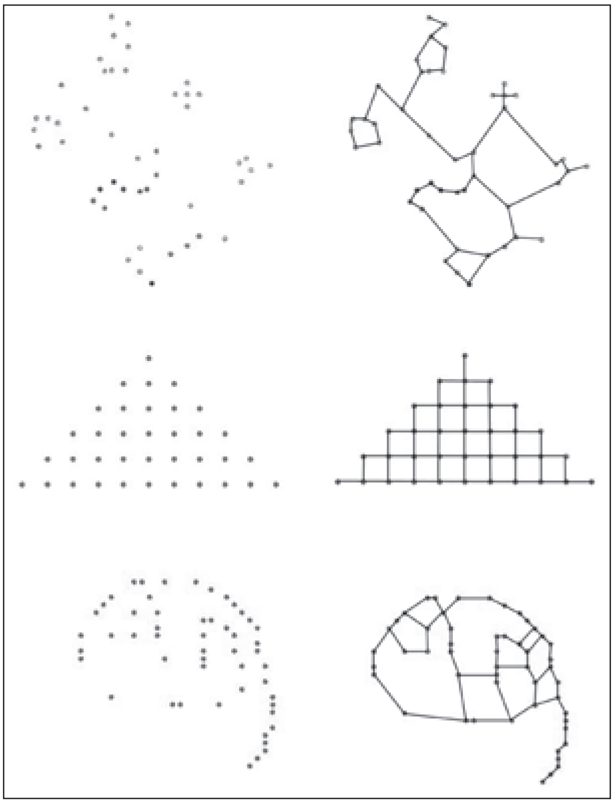


Figure 4. A sample of point sets and their corresponding Relative Neighborhood Graphs. Notice that the RNG captures the shape of the point distribution and therefore is a good descriptor of relationships between points.

The Gabriel graph is obtained by defining a circular region of influence instead of the almond-shape of the RNG (Figs 5 and 6).

Another interesting variation results when a parameter beta is introduced in order to enlarge or reduce voluntarily the region of influence for exploratory reasons. The parameter can be applied both to the almond-shaped region and the circular region mentioned above (Figs 7 and 8).

High values of Beta produce coarser networks, that is, structural views of point topology with few edges. On the contrary, low values of Beta produce networks with higher edge-density. Testing several values of Beta produces a parameterized family of networks known

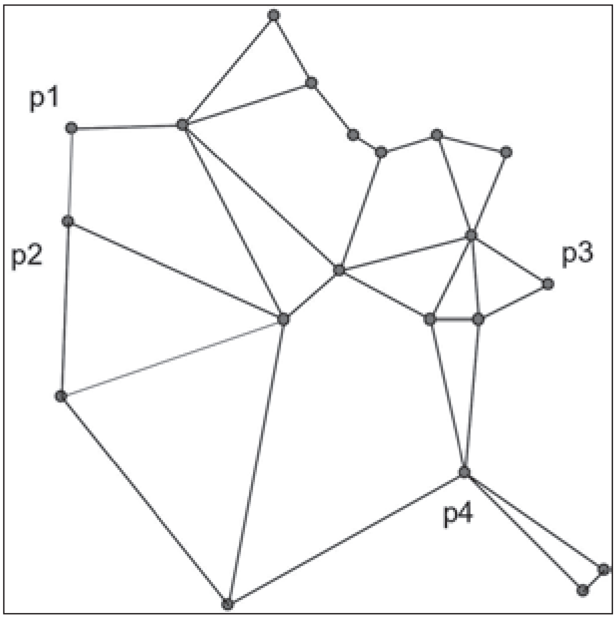


Figure 5. The test of "region emptiness" can also be applied using circular areas. The diameter of the circle depends on the separation of each pair of points.

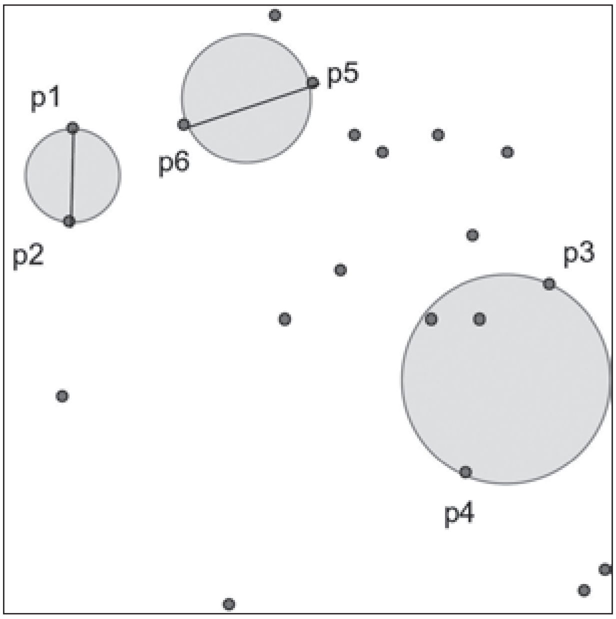


Figure 6. The so-called Gabriel Graph (GG). This is derived from applying the test of "region emptiness" using circular areas, instead of the almond-shaped regions of the RNG.

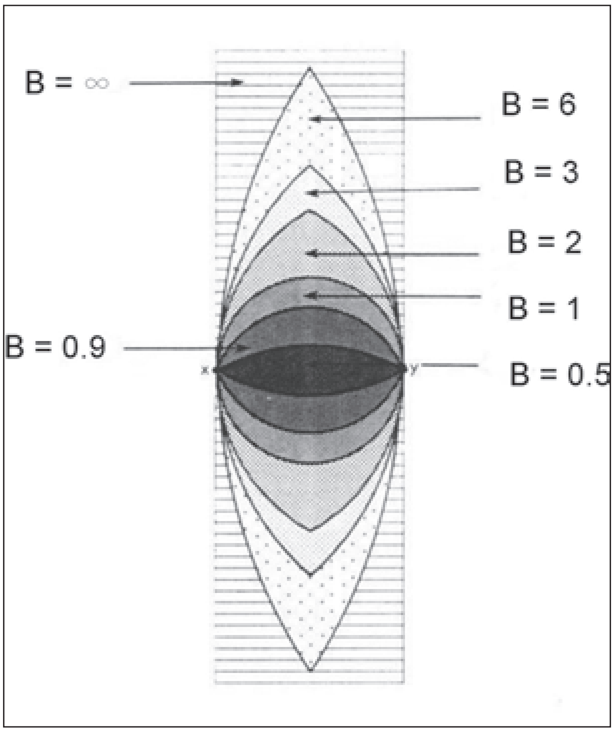


Figure 7. Applying a Beta parameter to the almond-shaped region of the RNG produces a family of graphs known as Beta-skeletons.

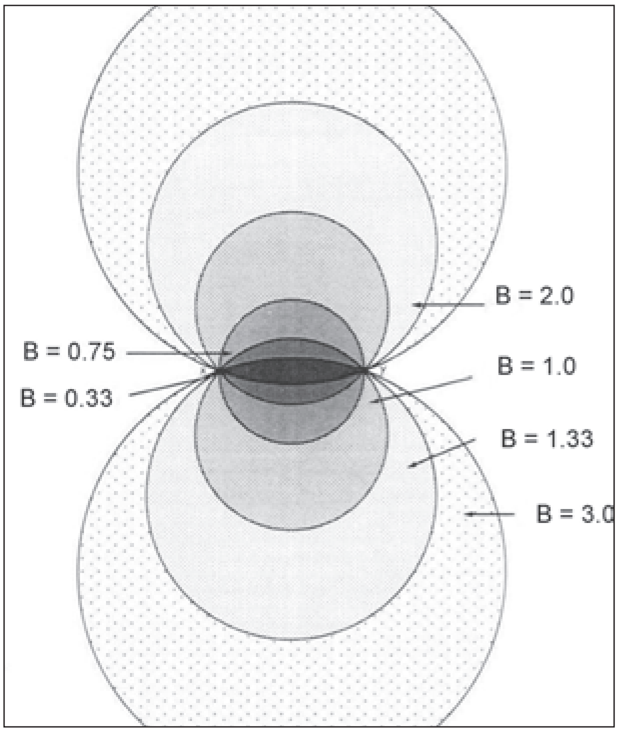


Figure 8. A second type of Beta-skeletons is produced by applying the parameter Beta to the circular regions of the GG.

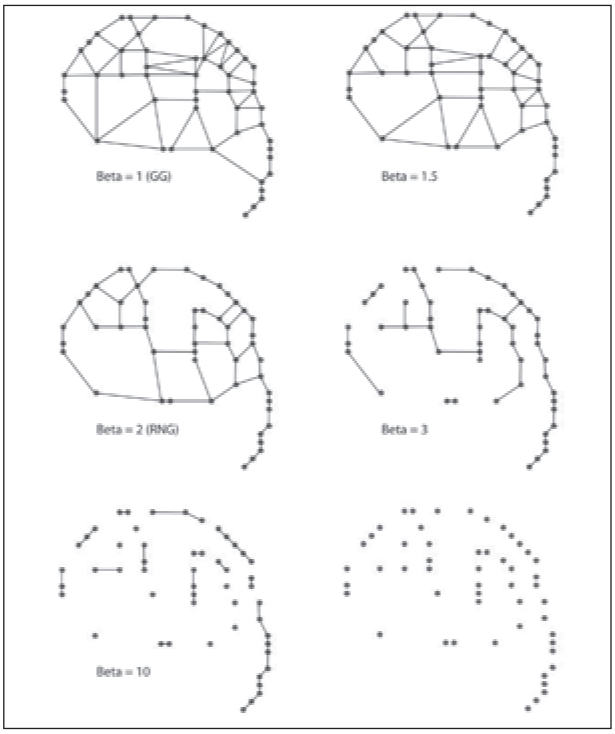


Figure 9. A point distribution and its corresponding family of Beta-skeletons. Among other things, these graphs can be used for exploratory applications in which it is necessary to analyze how strong or weak the links are.

as Beta-skeletons (Fig. 9). Two particularly interesting values of Beta are: 1) Lower Connectivity Threshold, defined as the highest value of Beta producing a connected Beta-skeleton with the lowest number of edges, that is the highest beta value before the network gets split into two or more sub-networks (after a certain value, most networks become disconnected), and 2) The second is Upper Connectivity Threshold, defined as the lowest Beta value that yields a connected Beta-skeleton with the largest number of edges. Given that they represent the minimum and maximum connectivity scenarios for a particular point set, either one of them can be used as a benchmark to perform inter-network comparisons (see more details in Jiménez-Badillo 2004, 166-172).

The goal of extracting Beta-skeletons is to obtain "...a spectrum of progressively more detailed descriptions of internal structure". In other words: "It allows us to visualize a spectrum of internal shapes of various edge densities. The entirety of this spectrum, including, in particular, the transitions between adjacent structures, provides an added dimension for the representation of structure. The second advantage is that it serves as a kind of benchmark with the aid of which empirical networks can be analyzed and, to some extent, compared" (Kirkpatrick and Radke 1985, 222).

Analysis of Relative Neighbourhood Networks

The advantage of the proximity networks described above is that they allow quantifying two important geometric properties, which can be used to derive archaeological conclusions. Inspired by the Space Syntax Theory developed by Hillier and Hanson (1984) in the field of architecture, and combining it with the notion of relative neighbourhood, we propose to focus on how symmetrical/asymmetrical a proximity network is and on how distributed/non-distributed its structure is. As we explain below, symmetry will lead us to discover patterns of integration/segregation, while distributedness will allow us to derive conclusions about the control structure of the network.

Symmetry

Given a proximity network, the relation of two nodes a and b is said to be symmetric if the relation from a to b is the same as the relation from b to a. The simplest case occurs when two nodes are directly adjacent (Fig. 10a). More complex examples include arrangements of three and more nodes. Figure 10b, for example, represents nodes a, b, and c having symmetrical relationships with each other. Observe the relation from a to b, which is the same as from b to a. The same happens from a to c and from c to a, as well as from b to c and from c to b.

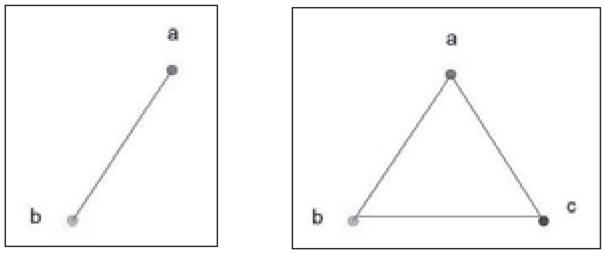


Figure 10. Two examples of symmetric relationships between points.

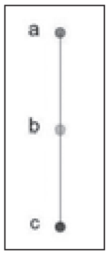
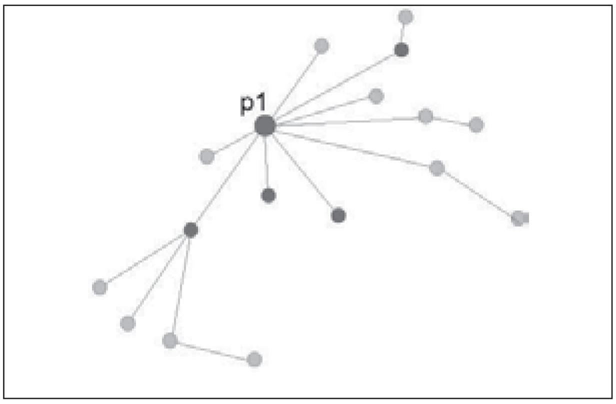
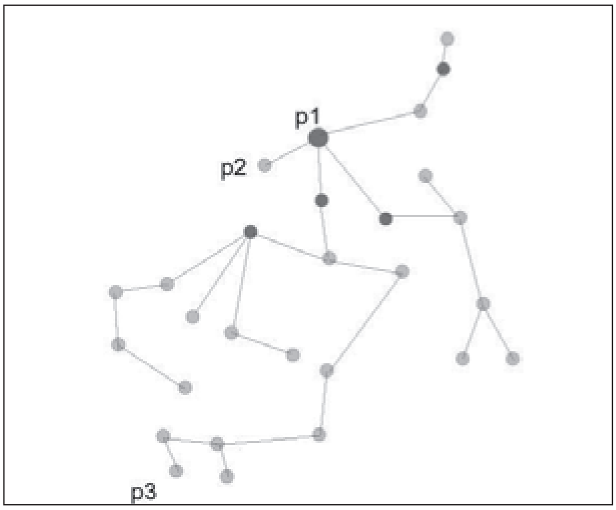


Figure 11. Asymmetric relationships between points.



12a



12b

Figure 12. Shallowness and depth resulting from two different network structures. In the example (a) the structure is shallow from node p1 because almost all the remaining nodes are only one edge away. In contrast, the structure shown in example (b) is deep from p1, because most vertices are many steps away. In the later case only p2 is one edge away, while the distance between p1 and p3 is 8 edges.

Asymmetry

In contrast, asymmetry occurs when one has to pass through an intermediate node when traveling from node a, for example, towards node c, as in figure 11. To reach node c from node a one necessarily has to pass through node b. Therefore, asymmetric relations necessarily involve a sense of depth, that is, step-distance from a certain starting-node to an ending-node. This depth is measured by counting how many edges exist in the path from the extreme nodes.

Measuring node integration

As it was said before, asymmetric relations involve the notion of depth or topological distance between nodes. A node has an asymmetric relation with regard to other nodes if it is located two or more steps away from them. In general, proximity networks where most nodes are a few steps away from each other are said to be *shallow*, while those networks whose nodes lay many steps away are said to be *deep* (Fig. 12).

Maximum shallowness exists when all nodes are connected directly to a single node. In this case, symmetric adjacency relations predominate among the nodes.

Maximum depth is registered in those systems where all nodes are arranged in a unilinear sequence away from one single node.

The value of integration for one particular node depends precisely on its shallowness or

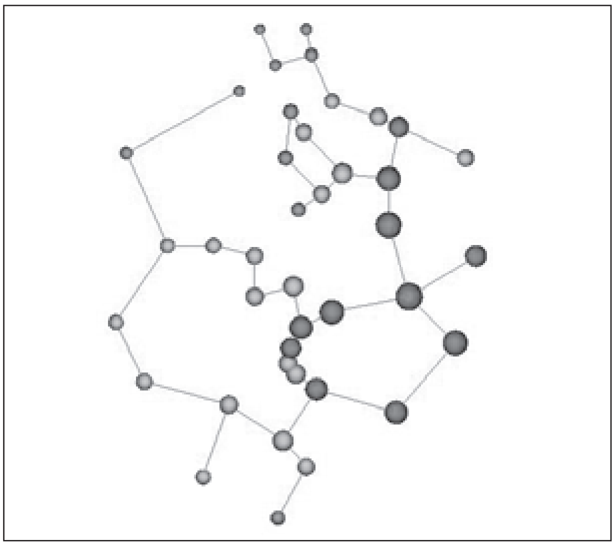


Figure 13. The integration structure of a particular network can be appreciated by scaling nodes according to integration values: the bigger the integration value the bigger the ball. In this example, the right side of the network is more integrated than the left, as shown by the red balls. The more segregated nodes are the small balls coloured in blue.

depthness with regard to all the remaining nodes of the network. In other words, integration measures the '*relative asymmetry*' from one particular node to all the remaining nodes of the network. The following procedure, adapted from Hillier and Hanson (1984, 109), allows extracting the value of relative asymmetry:

1. Given a proximity network G, take one node at a time, say p_i, and calculate depth values from that node to all the remaining nodes. Depth is just another name for the topological distance $td(p_i, p_j)$, obtained by counting how many edges are included in each path from p_i to p_j, for all p_j elements of P and p_j different than p_i.

$$Depth_i = \sum_{p_j \epsilon P; j \neq i} td(p_i, p_j)$$

2. Calculate the mean depth of p_i, denoted by md_i, by summing all the topological distances extracted previously (step 1) and dividing the result by the total number of nodes less one.

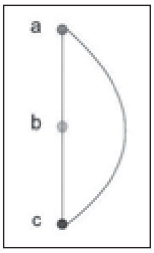


Figure 14. An example of distributedness between points.

3. Then, calculate the relative asymmetry RA_i of p_i as follows:

$$RA_i = \frac{2(md_i - 1)}{n - 2}$$

Where md_i is the mean depth of the ith node and n is the total number of nodes in the system.

4. Repeat operations 1 through 3 for every node of the proximity network. The procedure will give a value between 0 (maximum symmetry) and 1 (maximum asymmetry) for each node of the system. A value closer to zero indicates that the system is shallow from that particular node. In other words, symmetric relations predominate between that particular node and the rest of the system. For such reason, the node is well integrated into the system. In contrast, a value closer to 1 indicates that the system is deep from that particular node, asymmetric relations predominate and the node is segregated from the rest of the system (Hillier and Hanson 1984, 109).

Symmetry versus asymmetry equals integration versus segregation. Symmetric relations would reflect arrangements of artefacts or archaeological sites holding a certain amount of cohesion, unity, solidarity, etc. Alternately, asymmetric relations would identify those elements in a system characterised by isolation, separation, confinement, exclusion, etc. (Fig. 13).

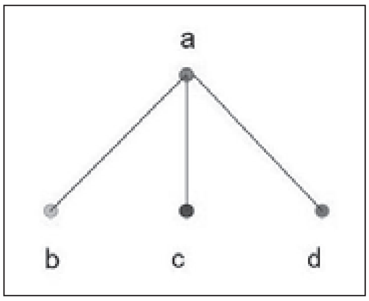


Figure 15. An example of non-distributedness between points.

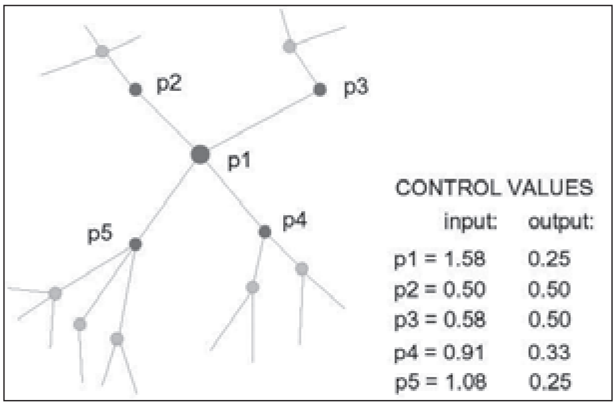


Figure 16. An example of vertex control. In this graph, p1 gives 0.25 of control to each of its four neighbors, while receiving 0.5 from p2, 0.5 from p3, 0.33 from p4, and 0.25 from p5. At the end, p1 acquires more control over the graph structure than its immediate neighbors.

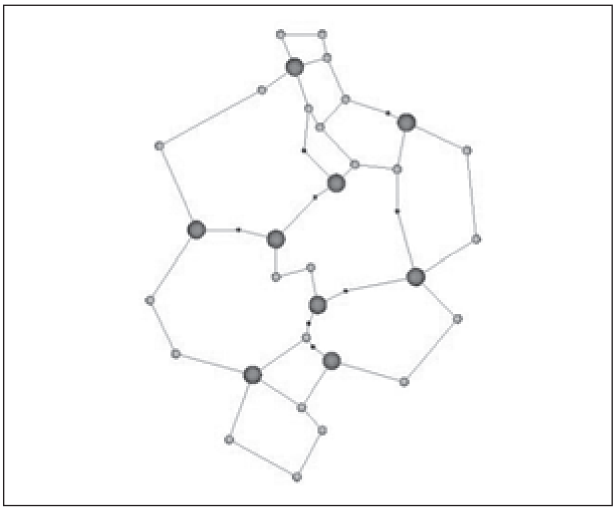


Figure 17. The control structure of a particular network can be apprciated by scaling nodes according to their control values. In this example, the big red balls represent nodes with bigger connectivity control than the small blue nodes.

Distributedness

A third property of proximity networks is called distributedness. This measures how many different paths are possible for traveling between two nodes given the whole connectivity structure of the network. In figure 14, going from node a towards node c presents at least two possible alternatives: the first one is a direct route between a and c; the second is the route passing through b.

Non-distributedness

Non-distributedness occurs when two nodes, say a and b, are related in such a way that there is one and only one path connecting a to b. In figure 15, for example, there is only one route to relate each pair of nodes and such path always passes through node a. This means that node a exercises some kind of control over the rest of the nodes. In fact, none of the nodes b, c, or d can communicate with each other unless they reach node a first.

Measuring node control

Distributedness yields a second measure called node control. This allows assessing the relative importance of artefacts or archaeological sites based on their local relationships.

Suppose that a certain node, say p_i, gives to each of its immediate neighbours $1/n_i$ of control (where n is the number of neighbours of p_i). At the same time, p_i receives a certain amount of control from its neighbours. This means that "... each space is partitioning one unit of [control] value among its neighbours and getting back a certain amount from its neighbours (Hillier and Hanson 1984, 109).

Nodes with values considerably greater than one exercise more control over the network structure than those nodes whose control value approaches zero. In the case illustrated in figure 16, for example, node p_1 partitions its unit of control among four direct neighbours (i.e. it gives 0.25 of control to each adjacent node), while receiving 0.50 from p_2, 0.5 from p_3, 0.33 from p_4, and 0.25 from p_5. The procedure exposes p_1 as the node of highest hierarchy with a total control of 1.58. In contrast, p_2 has the lowest control. It receives only 0.25 from each of its two neighbours, which gives a total control of 0.5. Therefore, distributedness versus non-distributedness equals control versus dependence. Distributedness would reflect cases of lower hierarchy, subordination, dependence, etc., while non-distributedness would reveal sites of higher hierarchy, control, power, authority, domination, etc. (Fig. 17).

Obviously, empirical archaeological systems such as sites within a region or artifacts in an archaeological context do not contain exclusively one type of relational pattern. On the contrary, they embody a finite number of combinations of symmetry, asymmetry, distributedness and non-distributedness. Hence, the task of the analyst is to measure how every node of an empirical system behaves in relation to those geometric properties.

To some researchers the computation time and effort for extracting the measures of control and integration of each node may seem excessive, especially when large datasets are involved. However, we have developed a computer program called Relative Neighbourhood Explorer that automatically and efficiently calculates the above measures and produces both the relative neighbourhood networks and the graph-profiles for each measure (for a user guide and explanation of the software see Jiménez-Badillo 2004).

Conclusions

There are two main advantages in incorporating proximity networks into archaeological spatial analysis, namely:

1. The notion of relative neighbourhood provides a mathematical formal way to determine whether one archaeological site or artefact has other significant spatial relationships besides its nearest neighbour. Therefore, the adoption of the concept represents an appropriate mechanism of exploratory heuristic analysis allowing us to discover contextual relationships, both at small and large scales.

2. The extraction of proximity networks not only provides a graphic view of interesting hypothetical connections. It also allows quantifying global and local spatial properties of the system. We proposed measuring the level of integration as a way to formally assess the degree of accessibility of certain nodes in relation to others. On the other hand, we proposed a control measure to rank the relative importance of particular archaeological sites or artefacts. This would be useful when sites appear to be of the same hierarchy at the beginning of the analysis and one needs to assess the impact of location given the specific spatial layout of the whole system.

As usual when dealing with this kind of approach, there are also some limitations. We start with points assumed to represent a coherent set. In archaeology such an assumption is not always justified because the set of archaeological sites may not be complete (some archaeological sites, for example, may not have been discovered at the moment of the study, some places may have not survived, or the chronology of the settlements may be unreliable). However, this limitation is not inherent to the proposed method. Rather, it is a problem related to the nature of the archaeological record. We must therefore treat any data set with the same caution exercised when applying other more traditional approaches.

On the other hand, the relative neighbourhood method is purely spatially oriented. It does not include any other variables in the analysis (though some tuning would make this possible). From this point of view it is limited, and therefore it must be seen as a starting point in the analysis of archaeological systems when we do not know the real links existing in the past. The reliance on space, however, could be an advantage in itself, as it does not impose previous assumptions on the data, that is, patterns of relationships are discovered rather than imposed.

Since 1998 we have been developing tools for spatial analysis, applying, in particular, proximity networks to study archaeological contexts such as the Mexica offerings (Jimenez and Chapman 2002, Jiménez-Badillo 2004, 2009a, 2009b).

In other countries, mainly the United Kingdom and Spain, other researchers have used the relative neighbourhood approach to answer archaeological questions. Particularly interesting is the work by Brughmans (2010) who has studied ceramic trade routes in the Mediterranean during the Roman Period using some of the proximity networks presented here. We expect that further developments would allow extending the range of applications for the proposed method to other kinds of problems.

References

Brughmans, T. 2010. "Connecting the Dots: Towards Archaeological Network Analysis." *Oxford Journal of Archaeology* 29 (3):277-303.

Gabriel, K. R., and Sokal, R. R. 1969. "A New Statistical Approach to Geographic Variation Analysis." *Systematic Zoology* 18:259-278.

Hillier, B., and Hanson, J. 1984. *The Social Logic of Space*. Cambridge: Cambridge University Press.

Huang, N.-F. 1990. "A Divide-and-Conquer Algorithm for Constructing Relative Neighbourhood Graph." *BIT* 30:196-206.

Hurtado, F., Liotta, G., and Meijer, H. 2001. "Optimal, Suboptimal, and Robust Algorithms for Proximity Graphs." *Lecture Notes in Computer Science* 2115:2-13.

Jaromczyk, J. W., and Toussaint, G. T. 1992. "Relative Neighbourhood Graphs and their Relatives." In *Proceedings of the IEEE* 80 (9):1502-1517.

Jaromczyk, J. W., and M. Kowaluk. 1987. "A Note on Relative Neighbourhood Graphs." In *Proceedings of the Third Annual Symposium on Computational Geometry*, 233-241. Waterloo, Canada.

Jaromczyk, J. W., and Kowaluk, M. 1991. "Constructing the Relative Neighbourhood Graph in 3-Dimensional Euclidean Space." *Discrete Applied Mathematics* 31:181-191.

Jimenez, D., and Chapman, D. 2002. "An Application of Proximity Graphs in Archaeological Spatial Analysis." In *Contemporary Themes in Archaeological Computing,* edited by D. Wheatley, G. Earl, and S. Poppy, 90-99. Oxford: University of Southampton Department of Archaeology Monograph 3.

Jiménez-Badillo, D. 2004. "A Method for Interactive Recognition of Three-dimensional Adjacency Patterns in Point Sets, Based on Relative Neighbourhood Graphs: An Archaeological Application." PhD diss., University of London.

Jiménez-Badillo, D. 2006. "El Análisis Espacial de las Ofrendas Mexicas por Medio de Gráficos de Proximidad." In *Arqueología e Historia del Centro de México: Homenaje a Eduardo Matos Moctezuma,* edited by L. López-Lujan, D. Carrasco, and L. Cué, 455-482. Mexico City: Instituto Nacional de Antropología e Historia.

Jiménez-Badillo, D. 2009a. "Una Aplicación de la Teoría de Gráficas en Arqueología (Primera parte)." *Carta Informativa: Revista de la Sociedad Matemática Mexicana* 60:1-8.

Jiménez-Badillo, D. 2009b. "Una Aplicación de la Teoría de Gráficas en Arqueología (Segunda parte)." *Carta Informativa: Revista de la Sociedad Matemática Mexicana* 61:1-8.

Kirkpatrick, D. G., and Radke, J. D. 1985. "A Framework for Computational Morphology." In *Computational Geometry*, edited by G. Toussaint, 217-248. North Holland: Elsevier Science Publishers.

Lankford, P. M. 1969. "Regionalization: Theory and Alternative Algorithms." *Geographical Analysis* 1: 196-212.

Matula, D. W., and Sokal, R. R.. 1980. "Properties of Gabriel Graphs Relevant to Geographic Variation Research and the Clustering of Points in the Plane." *Geographical Analysis* 12:205-222.

Nooy, W., Mrvar, A., and Batagelj, V. 2005. *Exploratory Social Network Analysis with Pajek*. Cambridge: Cambridge University Press.

Rao, S. V. 1998. *Some Studies on Beta-skeletons*. PhD diss. Indian Institute of Technology.

Su, T.-H., and Chang, R.-C. 1990. "The k-Gabriel Graphs and their Applications." In *Proceedings of the International Symposium SIGAL ´90,* 66-75.

Su, T.-H., and Chang, R.-C. 1991a. "Computing the k-Relative Neighbourhood Graphs in Euclidean Plane." *Pattern Recognition* 24 (3):231-239.

Su, T.-H., and Chang, R.-C. 1991b. "On Constructing Relative Neighbourhood Graphs in Euclidean k-Dimensional Spaces." *Computing* 46:121-130.

Toussaint, G. T. 1980a. "The Relative Neighbourhood Graph of a Finite Planar Set." *Pattern Recognition* 12:261-268.

Toussaint, G. T. 1980b. "Comment to the article: 'Algorithms for Computation of Relative Neighbourhood Graph' by R. B. Urquhart." *Electronic Letters* 16 (22):860-861.

Toussaint, G. T. 1980c. "Pattern Recognition and Geometric Complexity." In *Proceedings of the Fifth International Conference on Pattern Recognition,* 1324-1347.

Toussaint, G. T. 1988. "A Graph-theoretical Primal Sketch." In *Computational Morphology,* edited by G. T. Toussaint, 229-261. Amsterdam: Elsevier Science Publishers.

Toussaint, G. T., and Menard, R. 1980. "Fast Algorithms for Computing the Planar Relative Neighbourhood Graph." In *Proceedings of the Fifth Symposium of Operations Research*, 425-428, Kolhn.

Urquhart, R. 1980. "Algorithms for Computation of Relative Neighbourhood Graph." *Electronic Letters* 16 (14):556-557.

Urquhart, R. 1982. "Graph Theoretical Clustering Based on Limited Neighbourhood Sets." *Pattern Recognition* 15 (3):173-187.

Urquhart, R. 1983. "Some Properties of the Planar Euclidean Relative Neighbourhood Graph." *Pattern Recognition Letters* 16:317-322.