

Searching for 2D Spatial Network Holes

Femke Reitsma¹ and Shane Engel²

Geography Department, 2181 LeFrak Hall, University of Maryland College Park, MD 20742, ¹femke@geog.umd.edu, ²engeljs@umd.edu

Abstract. Research involving different forms of networks, such as internet networks, social networks, and cellular networks, has increasingly become an important field of study. From this work, a variety of different scaling laws have been discovered. However, these aspatial laws, stemming from graph theory, often do not apply to spatial networks. When searching for network holes, results from graph theory frequently do not correlate with 2D spatial holes that enforce planarity. We present a general approach for finding holes in a 2D spatial network, and in particular for a network representing street centrelines of an area in Washington, D.C. This methodology involves finding graph holes that can be restricted to 2D spatial holes by examining topological relationships between network features. These spatial network holes gain significance as the number of edges encompassing the hole, and the length of these edges increase. For this reason, our approach is designed to classify these holes into different sets based on the number of edges found and the length of those edges. The results of this application provide valuable insights in the nature of the network, highlighting areas that we know from experience are poorly connected and thus suffer from low accessibility.

Keywords: topological analysis, street network, urban morphology, network

1 Introduction

A network is a valuable representation of data. This data can be observably network in form, such as internet and street networks, or can be interpreted as a network, such as social and other relational networks. The benefit of representing data as a network lies in the graph theoretical measures available for network analysis. Such benefits have resulted in an explosion of research towards the development of new network measures, and the representation of old data newly dressed as a network [8, 7].

Instances like the New York power outage in mid August 2003, and a similar Italian crisis in late September 2003, have highlighted holes that exist in infrastructural networks. Similar holes can also be found in our social networks, giving power to certain

players, and limiting information flow to others [1]. Holes have also formed the topic of substantial philosophical debate [3]. The significance of holes lies in their ability to obstruct the flow of information or substance through or over a network. These areas tend to act as barriers to movement, where things are forced to travel around these holes in order to traverse the network. Whether that network represents social networks, transport networks, or any other data that can be represented as an adjacency matrix of relations, the hole forms a virtual object by affecting the rest of the network. Therefore, detecting such holes within networks provides us with information about parts of the network that are poorly connected. The automated discovery of 2D (two-dimensional) spatial network holes allows us to explore the impact of these holes on the 2D spatial network, thereby evaluating their importance to the processes operating upon this underlying fabric of structural relations [1].

This paper focuses on the algorithm for finding 2D spatial network holes, that is, those network holes that are found in planar-enforced topologies. This algorithm involves a number of steps, beginning with a search for graph holes. If these graph holes are established to remain holes under planar enforcement, they are graphically represented for visual analysis and measurement. This methodology is then applied to a case study of an urban street network in Washington D.C. The underlying thesis is that large holes are more important because they identify parts of the network that need to be connected. Hence, filling a hole is likely to be of more value for larger holes, however “larger” is defined [1].

The remainder of this paper is structured as follows; Section 2 discusses background issues regarding the search for holes. Section 3 describes the general methodology that forms the basis of a search for 2D spatial network holes, which is then applied to the case study of street networks in Section 4. Section 5 concludes the paper.

2 Background

Because space is inherent in the structure of spatial networks, these unique networks cannot be fully understood with common graphic theoretic measurements. Space confounds deterministic measures as it does stochastic measures through spatial autocorrelation. For example, research by [8] has found that scale-free patterns that are found in a large range of networks, such as the world wide web and social networks, do not exist in geographic networks, such as internet routers and power grid structures. Similarly, spatial holes do not conform to traditional metrics.

Loosely defined (see Section 3 for a formal definition), a network hole is a part of the network formed by four or more nodes that are connected by links, forming a non-repeating path where the first node equals the last. As a consequence of planar enforcement, graph holes do not equate to 2D spatial holes. In figure 1 below, the nodes (A, B, C, D) and the edges linking them form a graph hole. However, it does not form a spatial 2D hole as there is a crossing set of nodes and vertices, that is, nodes (D, E, B)

and the edges connecting them. Hence a general search for graph holes will include holes that are not 2D spatial holes.

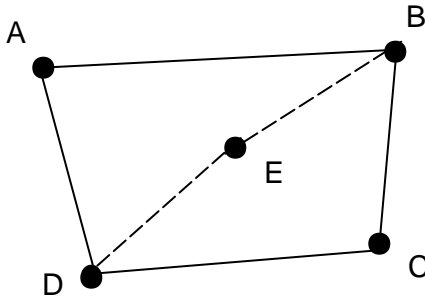


Fig. 1. Graph hole

Furthermore, the problem of detecting spatial network holes is not merely a problem of shortest path analysis from a node back to itself. As expressed in Figure 2, a spatially internal path (D, E, F, G, H, I, B) could be longer than the external path (A, B, C, D). In a shortest path analysis from a node such as A, this would result in (A, B, C, D) being selected.

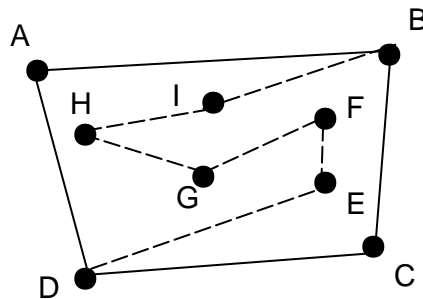


Fig. 2. Graph holes versus spatial holes

3 General Methodology

In overview, the methodology for finding 2D network holes involves three steps which couple GIS with a program implemented in Java. First, the network, represented as an adjacency matrix, is used to find graph holes. Second, the results of the graph hole search are spatially filtered for 2D spatial network holes. And finally, the output is visually represented for analysis.

3.1 Graph Hole Search

The problem of searching for holes in a network translates to a graph theoretical problem of searching for chordless cycles in a graph, also called graph holes. Given a set of vertices V and edges E , we form a graph $G(V, E)$. A graph cycle C is a subset of the set of graph edges that forms a path, a sequence of non-repeating edges, where the first edge is also the last. A chord of a cycle is an edge that joins two non-consecutive vertices of the cycle, that is, an edge not in the edge set of the cycle whose endpoints lie in the vertex set of C (Figure 3). A chordless cycle of graph G is a graph cycle of length at least four in G that has no cycle chord [4].

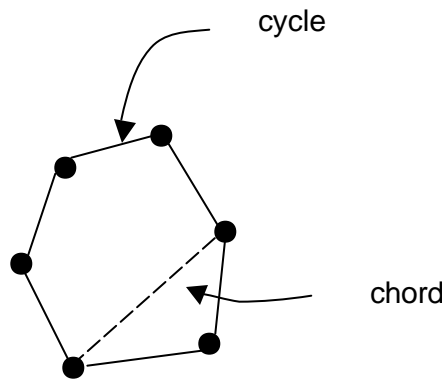


Fig. 3. Graph cycle and chord

The network to be analyzed must be represented as an adjacency matrix. An adjacency matrix is a matrix with rows and columns that are labeled with the network's nodes, or in terms of graph theory, graph vertices. The adjacency matrix $A = A(G) = [a_{ij}]$ is a $n \times n$ matrix in which the $a_{ij} = 1$ if v_i and v_j , where $v_i, v_j \in V$, are adjacent and $a_{ij} = 0$ otherwise. This adjacency matrix forms the input for the detection of chordless cycles or holes.

The method for detecting chordless cycles was informed by the work of [5]. It was implemented in the Java programming language using the JAMA matrix library (available online at: <http://math.nist.gov/javanumerics/jama/>). The algorithm is structured as follows:

```
for (each edge  $\{x, y\}$  of a matrix) do {  
    for (each chordless path  $\geq 4$ ) do {  
        if (the first vertex is connected to the last vertex) then{
```

```

        chordless path = chordless cycle
        write chordless path to array
    }
}
}

```

3.2 Spatial Hole Search

Searching for spatial holes is based on the output of the graph hole search. Each graph hole is checked to see whether it spatially contains a subgraph that connects at least two nodes of the hole. The JTS Topology Suite, which is an API of 2D spatial predicates and functions, was utilized for searching for the spatial holes (available online at <http://www.vividsolutions.com/jts/jtshome.htm>).

The general algorithm for searching for 2D spatial holes is as follows:

```

for (each graph hole) do {
    for (each node in the network) do {
        if (there is no node spatially within the hole that connects
            two unique nodes of hole) then {
            write hole to spatial holes array
        }
    }
}
}

```

The result of this algorithm is an array of 2D spatial network holes that can be used for visualization and analysis.

3.3 Spatial Hole Visualization

The array of 2D spatial network holes is transferred into ESRI's ArcGIS environment, and is manipulated with Visual Basic for Applications programming language. Each node from the vector street network is coordinated with its corresponding node in the array of spatial holes.

The basic algorithm for transferring the array of holes into graphic polygons is as follows:

```

for (each graph hole) do {
    for (each arc in the network) do {

```

```

    if (the edge contains two nodes that make up the hole) then{
        add edge vertices to point collection
    }
}
create new polygon from point collection
}

```

A resulting polygon geometry file is created to visualize the relationships between each hole. This output also provides the means to calculate some basic measurements, such as the area and perimeter of each hole.

4 Street Network Application

A subset street network from Washington D.C. was selected in order to test the algorithm and explore its potential in the analysis of street network structure. The street network is displayed below in Figure 4. It is already visually evident that parts of the network will be more difficult to traverse, in particular where there are large areas without any crossing streets.

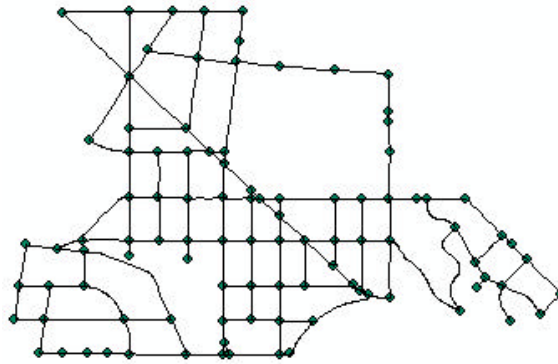


Fig. 4. Selected street network in Washington D.C.

4.1 Method and Results

The method implemented for finding 2D spatial network holes in street networks involved a number of steps beyond the general methodology described in Section 3 above. A street centerline network provided by the Washington, D.C. government was

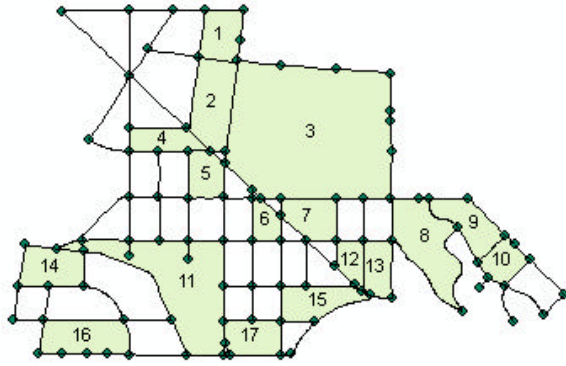
used for the analysis. Since these networks can vary in representation of real-world features, it is important to define some basic rules for the study area. The data represents the theoretical centerline of the actual ground feature; with no more than one edge for each route (divided highways and boulevards are often represented with more than one edge). Nodes define the start and end of each edge, and need to be placed where centerlines intersect each other. Overpasses need to be represented so that the edges that meet at this area do not intersect; in other words, without a node present at the intersection. For this example, private drives and alleyways were excluded, as this study is only concerned with main thoroughfares. Finally, any cul-de-sacs that are represented in a “lollipop” configuration should not be included, but should instead be shown with an ending edge.

As mentioned earlier, the vector street network is transferred into an adjacency matrix by describing the relationships between each pair of nodes in the network. Each node in this network has a relationship with one another through a connecting edge. From this understanding, it was necessary to use the general algorithm below to build the adjacency matrix:

```
create matrix consisting of columns and rows for each unique node
  for (each node in the network) {
    for (each edge in the network) {
      if (from-node or to-node of the edge matches the
          coordinates of node) {
        get the opposite node of either the from-node or to-
        node and write a 1 in the node row, opposite node
        column {x, y} value of the matrix
      }
    }
  }
}
```

Once holes were found from the adjacency matrix produced from the algorithm above, holes defined with four edges were removed from the array, as they do not greatly impact on accessibility of the network.

The results of the analysis are presented in Figure 5 below. The holes detected range in size from 5 edges to 15. Clearly hole number 3 provides the largest barrier to traffic crossing from one side of the street network surrounding the hole to another.



Hole	Area	Perimeter
1	23825	619
2	42004	890
3	255775	1983
4	19882	690
5	19676	556
6	12306	455
7	25280	637
8	53566	1182
9	27716	761
10	17713	520
11	105273	1794
12	19401	595
13	13898	501
14	30090	718
15	27720	796
16	36430	843
17	23173	618

Fig. 5. Results of 2D spatial network hole search (meters)

Table 1. Measurements of spatial network holes (meters)

Furthermore, we can see from examining the measurements in Table 1 that hole 3 also has the largest perimeter and area, correlating with the relatively large number of nodes that make up this region. We can therefore conclude that the network would benefit the most by having this hole crossed by an edge that divides the area and number of nodes the best. Unfortunately, the presence of the rather large and historic National Cathedral may make this slightly difficult.

4.2 Discussion

Finding holes within the street networks gains significance by providing an indication of barriers to accessibility. Street network holes can increase the cost of travel for individuals traversing the network from two locations on opposing sides of the hole. This is a purely structural approach to accessibility, defining the potential of the network for impeding the flow of traffic, rather than more traditional approaches which incorporate patterns of land use as activity elements [6]. It provides a measure of accessibility that is inherent in the network structure. Furthermore, holes identify areas that would be best suited for the addition of new roads, in order to benefit the entire network structure.

Street network holes cannot be inferred from other transportation network metrics as there are no graph theoretic measures that capture the spatial hole either directly or indirectly. The nodes and edges that form the hole do not give themselves away by local measures such as through lower node degree, that is, by fewer connections to those nodes

composing a hole. Nor can holes be detected through global measures such as the diameter of the network, which is the length of the shortest path between the most distant nodes in the graph, or indices such as the alpha index, which is a measure of connectivity comparing the number of cycles in a network with the maximum number of cycles.

4.3 Limitations and Extensions

Although this application is simple, it presents the first findings and analysis of spatial network holes. Yet there are limitations to the example presented in this paper, which directs us to extending this work. A limitation not recognized in example above may arise when the size of a hole in terms of the number of edges and nodes that describe it does not correlate with the area of the hole or perimeter of the hole. In other words, we could have a hole that has a large number of edges and nodes, but a rather small area and perimeter (or vice-versa). For example, roundabouts are prevalent in the D.C. area, resulting from the large number of streets that converge on these small areas, and can produce smaller holes that are densely populated with nodes. In these cases, naturally occurring in larger study areas, a ranking calculation would need to be developed incorporating the length of the edges in each hole.

Furthermore, in certain cases, holes defined with less than four edges could turn out to be large areas that act as greater barriers to the network than normal. An extreme scenario could present itself with the occurrence of a large area that is only connected through two edges, as in Figure 6. This is likely to be the case in mountainous terrain. This area would be excluded from the results of the methodology discussed above, as holes are defined as having equal to or more than four edges. A solution to this problem would be to increase the density of the nodes in the network by creating pseudo nodes if an edge exceeds a certain distance. These pseudo nodes would not define intersections but produce the same results in the analysis.

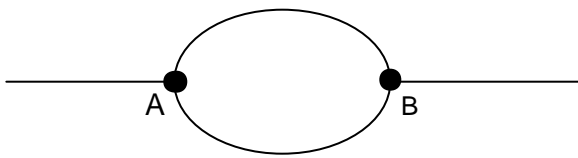


Fig. 6. Large hole consisting of only two nodes and edges

A logical addition would be to weight the links based on factors such as the number of lanes, usage, or other factors that can affect the impedance of a route. The results would then rank the holes according to importance based on these weights. For example, holes can be found at various levels of connectivity through slicing those levels at specific weights. In graph theory, the weighted adjacency matrix forms a graph with multiple

edges connecting vertices which can be represented as a network together with a function which maps the edge set into the set of integers. The resultant graph is termed a multigraph or hypergraph [2, 5]. Given a graph S is a subgraph of multigraph $M(V, E)$, where $V(S) \subseteq V(M)$ and $E(S) \subseteq E(M)$, the level of connectivity $i \in I$ of S is thus defined as the union of all edges with a value greater than or equal to i , and their connected vertices. That is, if

$$E_i \neq \emptyset$$

$$\bigcup_{i=I} E_i = E(S)$$

The general algorithm discussed above could then be applied at particular levels of connectivity by specifying the value to be found in the entry of the matrix.

Another approach to determining the importance of the hole is to analyze its impact on transportation paths. The hole could be weighted according to the number of times it impedes the accessibility of an individual trip. For example, a shortest path analysis could be executed for a specified origin-destination matrix, where if half of the nodes of a hole are used in the path, then that hole is impeding the flow. Such an approach integrates this analysis method to traditional measures typically based on Origin-Destination matrices. The hole structure could also be incorporated into other current metrics of accessibility, particularly gravity-based measures, where the hole may add to the impedance function [4].

5. Conclusion and Future Work

The search for 2D spatial network holes presents a new perspective on network structure, where the absence of network structure itself plays an important role in the nature of the network. Finding these holes allows us to evaluate the thesis that filling big and important holes. The application example of the street network provides some first insights into how these holes may impact on the network and its use for transportation.

Beyond the extensions proposed for street networks, this method has significant implications for networks in general, whether they are explicitly spatial or not. As expressed in the introduction, spatial networks, such as infrastructural networks, are susceptible to failure. An analysis of spatial holes may provide insights into the reasons behind these breakdowns. Furthermore, extending the dimensionality of these holes to three dimensions may provide insights into visual hollows that may have an impact on wayfinding research.

6. References

1. Atkin, R. H. (1977). *Combinatorial Connectivities in Social Systems: an application of simplicial complex structures to the study of large organizations*. Basel, Birkhauser Verlag.
2. Berge, C. (1970). *Graphs and Hypergraphs*. Amsterdam, North-Holland Publishing Company.
3. Casati, R. and A. C. Varzi (1994). *Holes and Other Superficialities*. Cambridge, MA, MIT Press.
4. Chandrasekharan, N., V. S. Lakshmanan and M. Medidi (1993). Efficient Parallel Algorithms for Finding Chordless Cycles in Graphs. *Parallel Processing Letters* 3(2): 165-170.
5. Chartrand, G. (1977). *Introductory Graph Theory*. New York, Dover Publications.
6. Handy, S. L. and D. A. Niemeier (1997). Measuring Accessibility: an Exploration of Issues and Alternatives. *Environment and Planning A* 29: 1175-1194.
7. Jeong, H., Z. Néda and A. L. Barabási (2003). Measuring Preferential Attachment for Evolving Networks. *Europhysics Letters* 61: 567-572.
8. Ravasz, E. and A. Barabasi (2003). Hierarchical Organization in Complex Networks. *Physical Review E* 67(026112).
9. Watts, D. J. and S. H. Strogatz (1998). Collective Dynamics of 'Small-World' Networks. *Nature* 393: 440-442.