ISSN 2616-5805 Vol. 5, Issue 3: 31-36, DOI: 10.25236/AJMS.2024.050305

Optimisation of the locations of emergency services in London using network science

Larry Cao

Harrow School, London, United Kingdom

Abstract: *The problem of how to optimise the locations of emergency services (e.g. hospitals, police stations, fire stations) around a city is a problem of utmost importance. It directly influences the well-being and the safety of civilians. Different arrangements of the locations of emergency service stations result in different lengths of time required for civilians to reach the nearest one for them. The purpose of this research paper is to produce the best distribution of the locations of emergency services through analysis using network science. The main hypothesis of this paper is that when considering multiple emergency service locations, the optimal distribution would be when the locations are placed away from each other as opposed to being clustered in the centre. The model included in this research paper utilises the London tube system for analysis and the Median Problem model for the method of analysis [1]. It outputs the optimum locations when considering one emergency service station, and when considering two emergency service stations. The results can be employed to help determine future construction plans for the government.*

Keywords: *Centrality; Networks; Nodes; Coefficient; Minimization*

1. Introduction

Network science finds diverse applications across many industries and businesses, serving various purposes. Some commonly seen applications may include social networks in different types of social media and market networks, where analysis is carried out on exchange relationships and price setting. In this paper, focus is put on how to optimally arrange emergency services around a city with a huge population-London. This research question could be beneficial to the wellbeing of many by minimising the danger that civilians face in their everyday lives. In terms of an application of network science that bears resemblance to the research question, the firm Flowminder provides a service which maps the optimal geographical locations of service stations so that the maximum number of people can be reached [2]. This firm employs a grid method to map the population distribution in a certain city, and it uses the service radius of the stations to maximise the percentage coverage of the civilians. Flowminder is already being used in many places, including Nigeria, where it maximises the number of children who have access to schools [2]. Also in Nigeria, Flowminder is used to place vaccination centres so that the largest number of people can get COVID19 vaccination [2]. This research paper can produce a similar effect in helping the society and minimising danger. The author will mainly be using quantitative data analysis and literature review to complete the research. Through the employment of databases and the incorporation of literature reviews, this paper will be able to find the optimal strategy and the suitable data to complete the researcher's network analysis. Then, through analysing and interpreting the numerical data the author will be able to optimise the locations of the emergency services around London. Computer simulations through Python are also present.

1.1 Introduction to networks

Networks have always been essential to mathematicians throughout the ages. It is an abstraction of the system in real-life that contains individual units called nodes and the relationships between nodes called edges. The main use of networks is quantification. Many properties of a system in real life, such as its structure and the importance of its individual nodes, can be expressed in terms of numbers when a network is created (Figure 1). A network also provides a visual representation of the connections between nodes and significant relationship properties such as clustering. One of the most famous and dated networks is the Map of Konigsberg created in Euler's time (Figure 2). Today, networks are used for countless different purposes, including social networks, geographical networks, and internet networks. Due to the simplified nature of networks, it does not take into account all of the immense number of

ISSN 2616-5805 Vol. 5, Issue 3: 31-36, DOI: 10.25236/AJMS.2024.050305

factors that play a role in shaping a system in real life. In addition, inaccuracies in data or missing data (noise) are often present when formulating a network.

```
g = nx.read_edgelist("London-2009-adjacency.txt",create_using=nx.Graph())
print(nx.info(g))
nx.draw(g)
print(nx.diameter(g))
N = g.number_of_nodes()
print(N)
```

Figure 1: Graph import and basic analysis

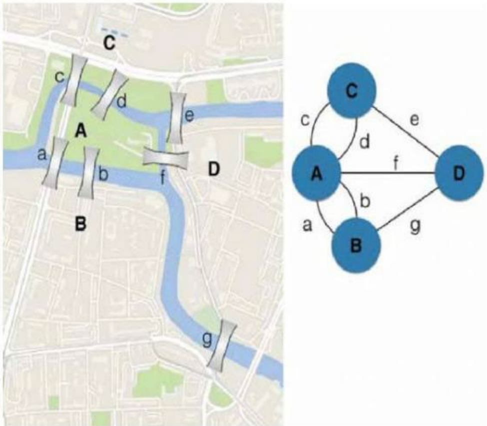

Figure 2: Map of Konigsberg, an abstraction of a geographical network created during Euler's time

1.2 Literature Review

A paper focusing on the topic of influence maximisation greatly impacted this research paper (Zareie and Sakellariou, 2023)[3]. The paper outlines the difference between behaviour-agnostic and behaviour-aware methods, and how both can be used to maximise the influence that a certain set of users/nodes have on the overall network. In addition, the paper examines how behavioural factors can be implemented into the network when exploring influence maximisation problems. It also considers the potential challenges that may arise during the modelling of an IM problem. This research paper takes inspiration from influence maximisation to see how a limited number of emergency services can be used in the optimal way.

Another piece of literature that influenced the author's research is a course on the topic of network science. This course explains the mathematics, applications, notations, and properties of networks. The researcher has implemented much of the paper, in particular the parts about closeness centrality, into this research paper. In addition, in the discussion section of the paper, the author has discussed the usefulness of other forms of centrality such as the Katz centrality and how they can be used to further improve the model. A large part of this research paper's method is based on the use of the Median Problem. The MIT course on this problem has been of massive help [1]. The author incorporated the method into real-world data to work out the optimal location when considering one node or multiple nodes.

1.3 Data

The data in this paper is obtained through the Colorado Index of Complex Networks, which is a large and diverse collection of helpful real-world networks [4]. It is one of the biggest tools mathematicians utilise for network analysis and research. In particular, the author made extensive use of the edgelist containing the positions of all the tube stations in London recorded in 2009. Much of the analysis and the method of this paper rely on data about the longitude and latitude values of the different tube stations, the distances between the tube stations, and how they connect with each other.

1.4 Summary of introduction

The results from the model suggest that it is beneficial to spread the locations out and establish emergency services in different parts of the city, supporting the main hypothesis. In particular, the model recommends one location to be in central London and one location to be in East London. More information about the results is included in the Results section below. The main takeaway from this research paper is the importance of quantitative analysis in research. Through quantitative analysis, we can arrive at accurate results, which may differ from intuition. Furthermore, this paper demonstrates the significance and the wide range of applications of network science. Future investigations into this field of mathematics will benefit countless other areas of study, such as neuroscience, computer science and sociology.

2. Results

2.1 Overview of method

To obtain the results, the researcher used the Median Problem model and Hakimi's theorem [1]. This research paper focuses on minimising the distances between emergency service locations and the individual nodes throughout the city. Full details on the method are included in the Materials and Methods section below.

2.2 One location

The researcher created a colour-coded network showing all the nodes (Figure 3, Figure 4) to depict the tube stations with the highest closeness centralities in a clear manner. Through NetworkX, the code returned the five stations with the highest closeness centralities. The tube station with the highest closeness centrality was "Green Park" (Figure 5). The full copy of the code is included in the appendix section of the paper.

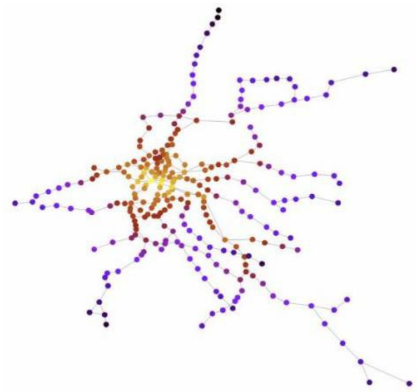

Figure 3: Colour-coded network based on closeness centralities of nodes

```
pos_latlong_london = {}
for i in g.nodes():
    pos_latlong_london[i] = [g.nodes[i]['long'], g.nodes[i]['lat']]
```

```
plt.figure(figsize=(20, 20))
nx.draw_networkx_edges(g, pos_latlong_london, alpha=0.4)

edges = nx.draw_networkx_nodes(
    g,pos_latlong_london,
    node_size=150,
    cmap=plt.cm.gnuplot,
    node_color = a)

plt.axis("off")
plt.show()
```

Figure 4: Code for Figure 2

ISSN 2616-5805 Vol. 5, Issue 3: 31-36, DOI: 10.25236/AJMS.2024.050305

```
sorted(clo_cen, key=clo_cen.get, reverse=True)[:5]

['GreenPark', 'BondStreet', 'OxfordCircus', 'BakerStreet', 'Westminster']
```

Figure 5: Tube stations with the highest closeness centralities

2.3 Two locations

After exhausting every combination, the two optimum tube stations for placing emergency services were determined to be "Stratford" and "Bond Street". When every node was optimally assigned to one of the two locations, the sum of all the shortest distances added together had a value of 1958 (Figure 6).

```
Stratford
GreenPark
1958
```

Figure 6: Two tube stations with lowest total distance

3. Discussion

3.1 Analysis of results

3.1.1 One location

The result for one emergency location is accurate when considering the locations of the tube stations in real life. The researcher has put all the focus on finding the nodes with the highest closeness centralities, meaning that improvements could be made by adding more emergency locations into the model.

3.1.2 Two locations

The result for two locations is reasonable, as shown by the locations of Stratford and Bond Street in real life. Stratford is in East London while Bond Street is situated in the centre. A large area can be covered with these two tube stations as the locations of the emergency services. The model can be improved by returning more than two tube stations. Realistically, the resources owned by the government can support many emergency services around the city.

3.2 Limitations

The researcher's network analysis did not account for several real-world complexities. Limitations exist because the real-life system is extremely complicated, and not enough computing power is available to consider every factor. Inevitably, limitations lead to inaccuracies in the results.

3.2.1 General Limitations

The edges in the network may not represent the exact real-life distances between two stations. Additionally, population density is assumed to be constant across London, which is not the case in reality.

3.2.2 Two Locations

Hakimi's Theorem suggests that a set of k-medians does exist strictly on the nodes [1]. However, due to the massive scale of the London Underground System (266 nodes in the 2009 network), it would be extremely difficult for the researcher to include every single node into distance calculations. Therefore, although the locations of the 10 nodes selected are representative of the whole city, the actual optimum locations are unlikely to be all included in the 10 nodes. For a more accurate model, more calculating power is required to run every combination of multiple tube stations out of the potential 266 candidates.

ISSN 2616-5805 Vol. 5, Issue 3: 31-36, DOI: 10.25236/AJMS.2024.050305

3.3 Potential improvements

In this model, the author could have incorporated different types of centralities such as the Katz centrality and the betweenness centrality in different cases to further improve accuracy. Betweenness centrality measures the node's relevance when considering shortest paths. The Katz centrality is measured using a node's immediate neighbours and the other nodes connected to that node through its neighbours, but the further the connection, the smaller the weight. In addition, this research paper could be improved by incorporating data on the population density distribution around London into the model. This would allow more weight to be put on areas with more people, which means more danger and emergencies.

4. Materials and methods

4.1 One location

Firstly, the researcher considered the case where only one emergency service station was available for the whole city. Through analysis using NetworkX and python, the node with the highest closeness centrality was identified (Figure 7). The closeness centrality of a node v_i in a network is defined as $\frac{N-1}{\sum_{j=1,j\neq i}^{N} d(v_i,v_j)}$. A useful model for considering the optimal locations of one or multiple emergency stations is the median problem [1].

```
clo_cen = nx.closeness_centrality(g)
a=[]
for i in clo_cen:
    a.append(clo_cen[i])
```

Figure 7: Creating a list of closeness centralities for the nodes

4.2 Median problem mathematical model for one location

Consider an undirected network $G(N, A)$ with n nodes. Let k be a positive integer ($k = 1,2,3, \ldots$) and choose k distinct points on the graph G to be represented as the set $X_k = x_1, x_2, \ldots, x_{k-1}, x_k$. Then, represent by $d(X_k, j)$ the minimum distance between any one of the points $x_i \in X_k$ and the node j in the network G.

$$d(X_k, j) = \mathrm{Min}_{x_i \in X_k}\, d(x_i, j)$$

The k-medians of network G is defined as a set of k points X_k^* such that, for every $X_k \in G$,

$$J(X_k^*) \leq J(X_k)$$

where

$$J(X_k) = \sum_{j=1}^{n} h_j d(X_k, j)$$

For a single node location problem, the median problem and the closeness centrality formula are logically identical. Both aim to identify the optimal node position by minimising the sum of distances to all other nodes - the median problem does so directly, while closeness centrality inverts and maximises the reciprocal of this sum. Therefore, both measures identify the same node as being the optimal location in the network structure.

4.3 Two locations

To explore the scenario involving multiple nodes for emergency service locations, a fundamental understanding of Hakimi's Theorem is crucial [1]. Hakimi's Theorem states that at least one set of k-median exists strictly on the nodes of G. This theorem is essential to the analysis of this problem by demonstrating the existence and discoverability of solution sets involving multiple nodes. The new

Median Model uses a shortest-distance matrix, which includes the values of the shortest distances between all pairs of nodes, i and j, of the network. To analyse the placement of two emergency locations, the same matrix can be used. Every combination of node pairs must be considered. When the node pair is confirmed, demands from other nodes are automatically assigned to the node with the lowest value of $d(i, j)$.

This research paper reduces the complexity of the problem by choosing the most important nodes in each part of the city instead of including every single tube station. Additional information regarding this approach can be found in the Limitations section of the paper. Through NetworkX, a more sophisticated result of two emergency locations out of ten candidate locations can be obtained. This process utilises code that systematically explores all possible combinations of two nodes from a pool of ten candidates for emergency services (Figure 8). The code calculates the sum of the shortest distances, considering the assignment of each individual node in the network to its nearest emergency location.

```
sum=0

station_options = ["Brixton", "Wimbledon", "Stratford", "WestHam", "WembleyPark", "ManorHouse", "Hampstead", "EalingBroadway",
                   "GreenPark", "BondStreet"]

for j in station_options:
    for k in station_options:
        for i in g.nodes():
            a=nx.shortest_path_length(g, source=i, target=j)
            b=nx.shortest_path_length(g, source=i, target=k)
            c=min(a,b)
            sum=sum+c
        print (sum)
        print (j)
        print (k)
        sum = 0
```

Figure 8: Code for analysing multiple tube stations

5. Conclusion

In conclusion, the results obtained by the researcher authenticates the hypothesis that an optimal distribution of emergency services occurs when the locations of emergency service stations are spread out in different parts of the city. The model offers a glimpse of what the basis of a potentially useful distribution strategy looks like. However, consideration of other factors, such as varying population density, would make the model much more comprehensive and practical.

References

[1] "Section 6-5-2.", self-contained, MIT, Web.mit.edu, web.mit.edu/urban_or_book/ www/book/ chapter6/ 6.5.2.html
[2] "Service Site Placement Optimisation." Www.flowminder.org, www.flowminder.org/what-we-do/service-site-placement-optimisation, Accessed Jan. 2024.
[3] Zareie, Ahmad, and Rizos Sakellariou. "Influence Maximization in Social Networks: A Survey of Behaviour-Aware Methods." Social Network Analysis and Mining, vol. 13, no. 1, 25 Apr. 2023
[4] "Index of Complex Networks." Index of Complex Networks, https://icon.colorado.edu, Accessed Jan. 2024