Internship Report

Finding shortest path covering maximum stoppages in a road network

Submitted by: Trisha Kanji, B.E. III year, Computer Science and Technology, IIEST, Shibpur

Start Date for Internship: 12th May 2016
End Date for Internship: 30th June 2016

- **Introduction**

  The report presents the following tasks completed during the summer internship at IIT Kharagpur:

  - Representation of a set of coordinates (here, bus stoppages) mapped on the map of Kolkata city forming a route as a graph structure.
  - Finding the shortest path to reach a stoppage from the previous one as mentioned in the bus route. In this case, the bus route covers all the stoppages in minimum time.
  - Next, when one or more stoppages become inaccessible, finding a shortest possible route avoiding those stoppages such that maximum number of other stoppages is covered and time taken is minimum.

- **Methodology**

  Data used consists of present day bus routes of Kolkata city. Each bus route consists of a set of stoppages. The coordinates of the stoppages are mapped to the road map of Kolkata city, which thus represent a graph.

  All the roads are represented as set of nodes and edges. To find the shortest path between two stoppages we map the coordinates of both on the

road map of Kolkata. The near most nodes on the graph for both are taken as the source and the destination nodes respectively. Next shortest path is obtained using Dijkstra's algorithm.

The final task consists of presenting an effective and optimal route in case of disruption of one or more stoppages in the present bus route. The approach used consists of two factors:

1. The number of stoppages excluding the disrupted ones should be maximized.
2. The time taken should be minimum. Here, the assumption made is that the velocities with which the buses travel are constant, which makes the time directly proportional to the distance covered.

Mathematically representing these two factors leads to an expression described below, maximizing which helps in achieving the result,

$$\sum_{i=1}^{n}[stopped(ST[i]) - (\frac{1}{n} * \left|d_{ST[i]} - \delta_{ST[i]}\right| * stopped(ST[i]))]$$

where,
n= number of stoppages
ST[]=stoppages (coordinates)
$\delta_{ST[i]}$ = shortest distance from ST[i-1] to ST[i] in initial route/ graph
$d_{ST[i]}$ = shortest distance from ST[i-1] to ST[i] in the route/ graph (after removing the disrupted stoppages)
Stopped()➔{0,1}, where 0-false, 1-true

- **Experiments and Results**

  The snapshots of results obtained for a few routes are shown.

Fig.1 Stoppages in purple and path in green circles

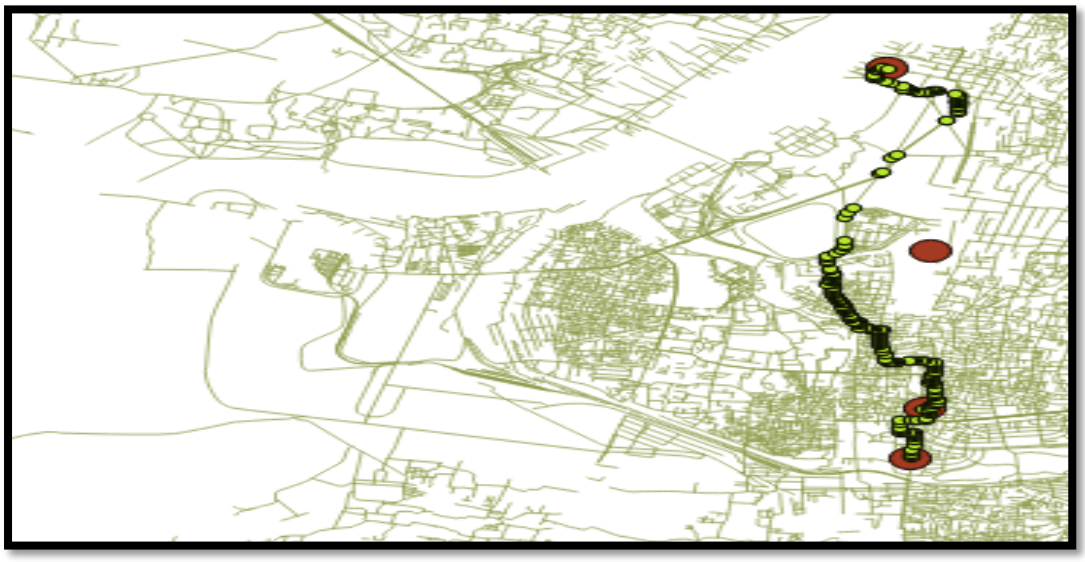

Fig.1b Stoppages in maroon and path in green circles

Fig.1a shows the initial route with 4 stoppages while Fig.1b depicts the alterative path when one of the 4 stoppages becomes inaccessible. Similar for Fig.2a and Fig.2b.

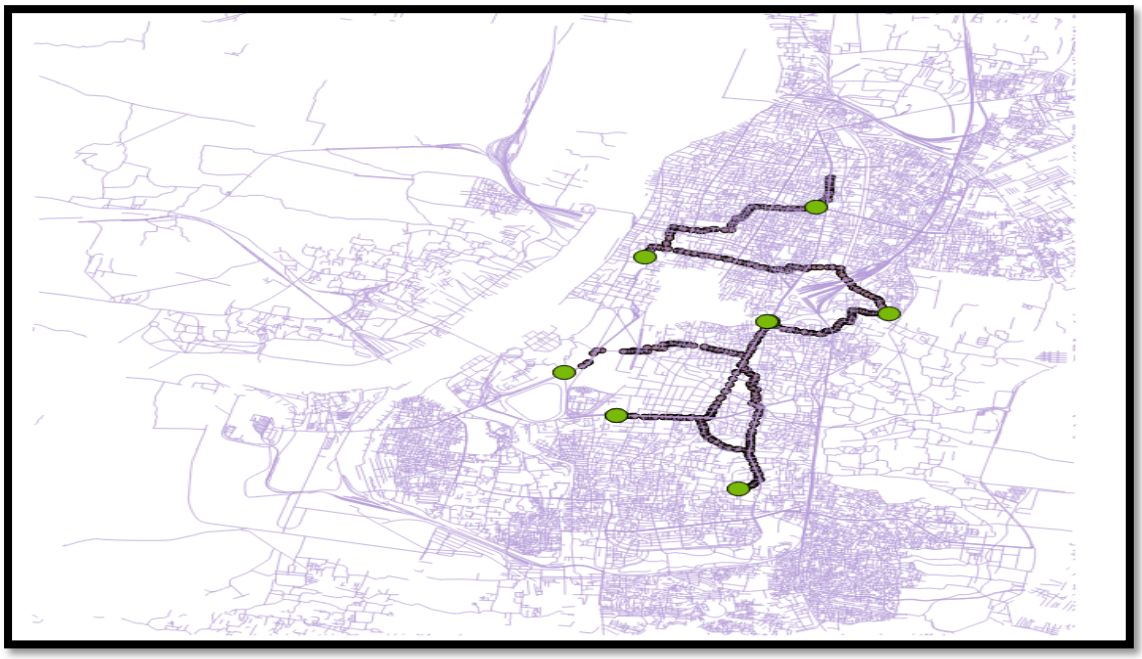

Fig.2a Stoppages in green and path in green circle

Fig.2b Bigger circles represent stoppages and smaller ones the path

- **Problems faced**

  - A number of stoppage points (coordinates) mapped on to the road map did not exactly fall on the road network. This required those points to be matched to the nearest roads.(refer Fig.3a and Fig.3b)

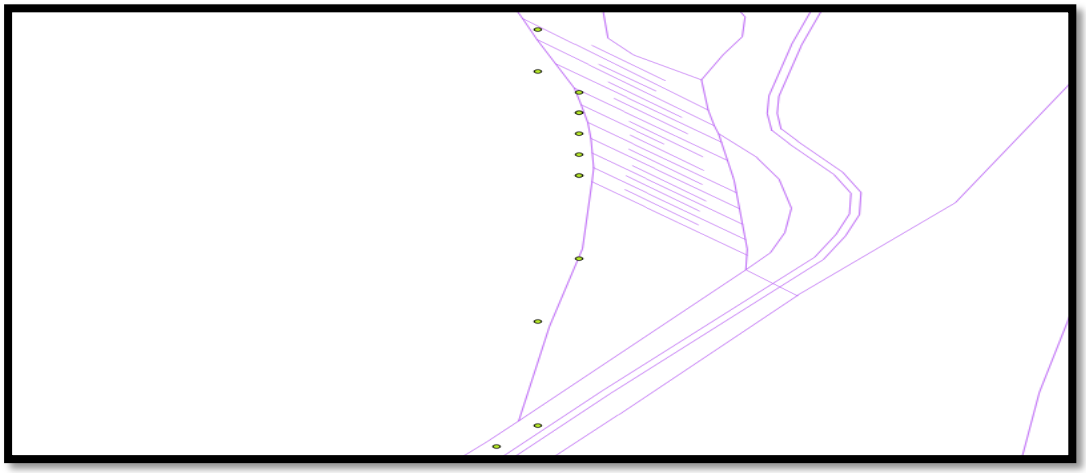

Fig.3a Points (in red) mapped to the road network (a number of which do not fall on the roads)

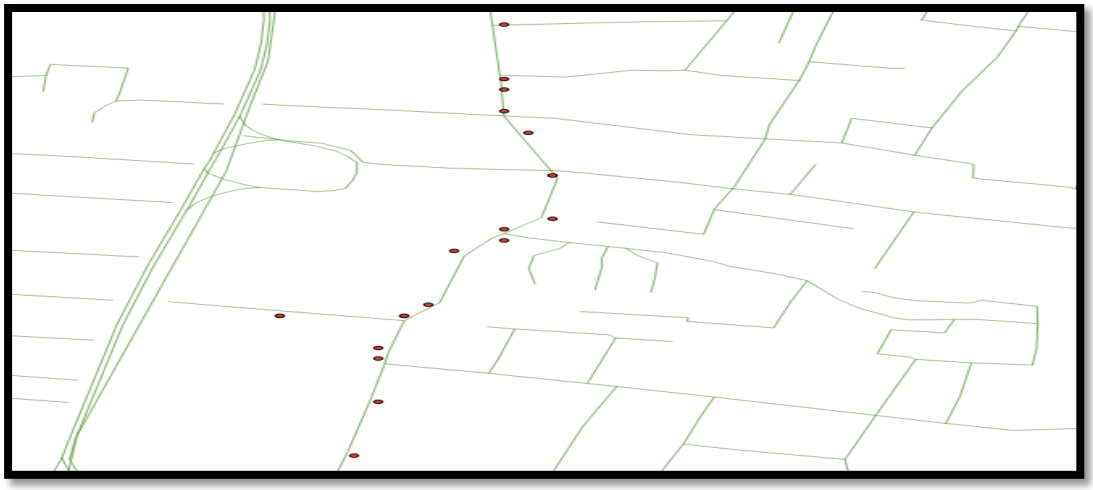

Fig.3b Points (in red) mapped to the road network (a number of which do not fall on the roads)

- The graph obtained from the given road network had a lot of disconnections due to the presence of innumerable disconnected small and irrelevant roads (such as roads inside parks, residential complexes, footpaths etc.). Such roads need to be cleared out from the map. (refer Fig.4a and Fig.4b)

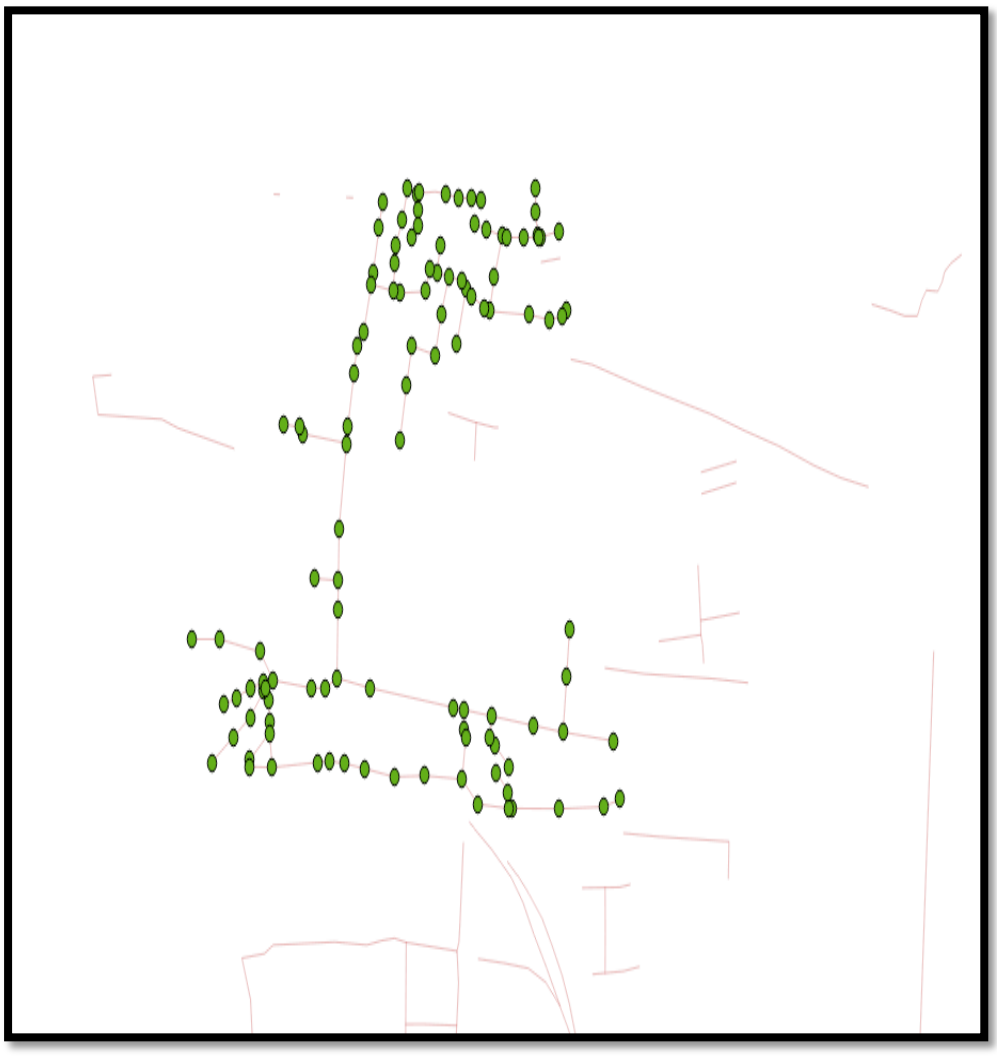

Fig.4a A disconnected graph component on the Kolkata road network (nodes in green)

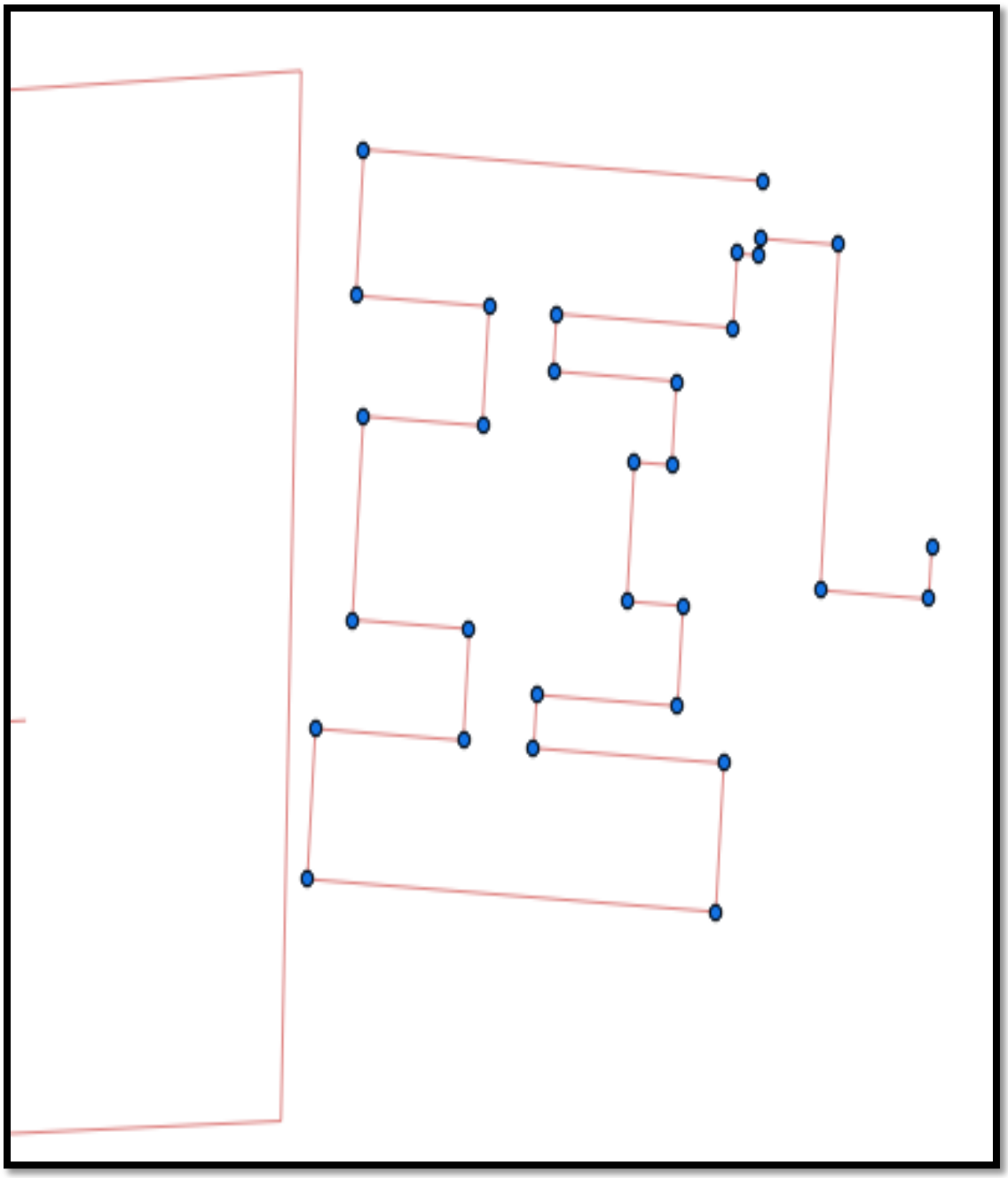

Fig.4b A disconnected graph component on the Kolkata road network (nodes in blue)

- **Conclusion**

  While working on this project I had to use and implement a lot of concepts in the subjects of data structure and algorithms, database management systems, discrete structures and computer networks which I have learned in the past few semesters in my coursework. The real life implementation of these concepts made me realize two very important things, first that real life data is huge, complex and different from what we expect it to be, and second that no single algorithm is perfect for implementing a real life task, there is always a cluster of algorithms for doing so. Finally I would like to state that the work that I have done here as an intern has given me some good insight into real life implementations and datasets, the various problems faced while dealing with them and handling them. Apart from the theoretical concepts in the coursework, these first hand experiences are something that I am going to require to use for sure for future work purposes.