2016 International Symposium on Nonlinear Theory and Its Applications,
NOLTA2016, Yugawara, Japan, November 27th-30th, 2016

Bus Transport Network in Hong Kong: Scale-free or not?

Tanuja Shanmukhappa, Ivan Wang-Hei Ho, and Chi K. Tse

Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong
The Hong Kong Polytechnic University Shenzhen Research Institute, China

Email: tanuja.shanmukhappa@connect.polyu.hk, ivanwh.ho@polyu.edu.hk, encktse@polyu.edu.hk

Abstract– Bus Transport Network (BTN) is one of the major sub-networks in Public Transport Network (PTN) in any city for daily commute of a large number of passengers. In this paper, we analyze the BTN structure in Hong Kong considering the geographical locations of bus stops and their directed connectivity from a complex network perspective. The presence of various geographical and socio-economic constraints in the city give rise to the bus network structure of 4,065 nodes connected with 7,909 edges. 916 routes are operated by five major franchised bus services in Hong Kong. Through introducing and implementing a novel 'Supernode' concept in the BTN, which is a collection of closely associated bus stops with respect to their separations and directions, our work evidently shows that the unweighted, directed network under consideration behaves more closely to a scale-free network after applying such concept. Such finding lays the foundation for future studies on BTN design and bus route management for optimal transport efficiency.

1. Introduction

Being one of the most densely populated cities in the world, one third of the total daily public transport in Hong Kong (HK) is accounted by the Bus Transport Network (BTN) according to the latest census statistics [1]. To facilitate the design and management of the bus system for optimal transport efficiency, the understanding of the topological properties, network structures and the connectivity associated with the bus network is of crucial importance. In addition, the evolution of the city has a major influence on the development of the BTN in HK. For example, before the opening of the Cross-Harbour Tunnel connecting Kowloon and the Hong Kong Island in 1972, China Motor Bus Co. (on the Island side) and Kowloon Motor Bus Co. (on Kowloon side) were the major bus operators. Nowadays, Hong Kong has five major franchised bus services operating throughout the city.

In this paper, we aim to analyze the BTN in HK from a complex network perspective. Network analysis started gaining more importance when the Erdos-Renyi method was first proposed for generating random graphs [2]. However, it was evidently recognized later on that the topologies and evolution of real world networks in our daily lives are governed by much more advanced concepts in the field of complex networks. Complex networks can be exploited to describe a wide range of systems from PTN to the Internet, financial systems to social networks, etc. Its applications cover many fields of systems in the real world [3].

There are a number of previous studies on bus network analysis. For example, a detailed study of the bus network structure and its statistical analysis of five different cities in India was provided by Chatterjee *et al.* in [4]. This paper also discussed a similar concept in brief called the short-distance station pairs, which combines stops that are geographically close to each other (walkable within few meters) with no direct bus connectivity. Zhang *et al.* [5] analyzed the BTN in Beijing based on complex networks defined in the Space-L and P concepts, where Space-L was used for the analysis of topological properties and Space-P for the transfer properties. A weighted complex network analysis of the travel routes in Singapore was done by Soh *et al.* [6], in which the topological and dynamical properties of the graph structure were considered for analyzing the transport network. This was the first paper that discussed the study of PTN based on geographical properties (e.g., latitude and longitude information). Ferber *et al.* [7] modeled and empirically analyzed PTN in 14 different cities across the world by a systematic approach under a number of graph representations. A more exhaustive work in the domain was done by Sienkiewicz *et al.* [8] on the statistical analysis of 22 public transport networks in Poland using various concepts of complex network.

With respect to the network size considered in all of the afore-mentioned works, it is evident that Hong Kong gets one of the most densely connected network structures with more than 4,000 nodes, 7,000 edges and 916 routes in total for all the franchised bus services. Our work here aims at analyzing the geographically-constrained bus network in Hong Kong using complex network concepts. The dataset is extracted from the centralized database from the Hong Kong Government [9]. Considering the bus stops and route connectivity among all the operators, a directed graph is generated in the L-Space based on the connectivity list in the dataset. The network is verified for the scale-free property as per the Barabasi-Albert model [10]. Initially, the network does not show strong scale-free property with node representation according to the standard graph theory, but with the newly introduced concept of supernode, we found that the structure behaves more like a scale-free network. Such finding will provide insights into the studies and design of bus network and its management. For instance, bus network interpreted with the supernode concept can be easily compared and evaluated with other theoretical scale-free networks generated according to the BA model. Hence, better network design can be identified and management decisions can be made for optimal transport efficiency.

2. The Bus Transport Network in Hong Kong

2.1. Graph Structure

The dataset collected from the centralized repository of the Hong Kong Government [9] is used for the generation of the directed, unweighted graph for the BTN in HK. The graph has n nodes, where each node represents a bus stop, and e edges which are connected in Space-L (there exists an edge between two nodes if there is a bus route connecting them directly). Hence the graph is a collection of n nodes and e edges: $G= (N, E)$. Where the set $\boldsymbol{N} = \{N_j, j=1,2...n\}$, set $\boldsymbol{E}=\{E_k, E_k=(N_i,N_j); N_i,N_j \in N \;\&\; k=1,2.....e\}$. In the digraph, N_i denotes the edge tail while N_j denotes the edge head.

A n x n adjacency matrix is used to describe the connectivity in G with entries $\{a_{ij}\}$. Where $a_{ij} = 1$ if there exists a link between N_i and N_j, and 0 otherwise. The adjacency matrix is then converted to the edge list which is input to the graph generation using the *Gephi* tool.

In our case, every node is identified by a unique Node_ID (NID), Node_Name (NN), Latitude (Lat) and Longitude (Long) information. The Lat-Long information is based on the 'WGS84' datum standard [11]. Fig. 1 shows the complete distribution of nodes (bus stops) in Hong Kong as viewed on Google Earth based on their Lat-Long information.

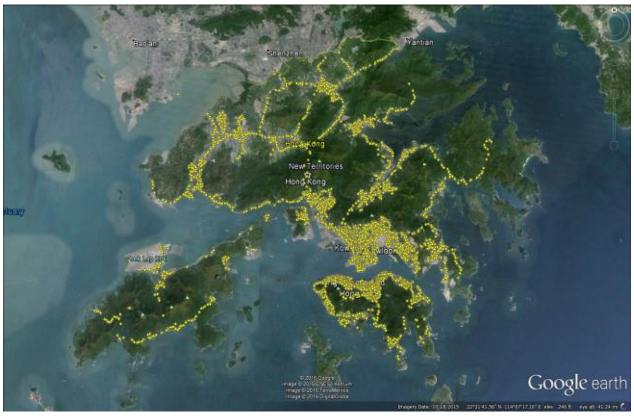

Fig. 1. Bus-stop distribution in HK based on Lat-Long.

The node locations are fixed based on the geographical coordinates and the connectivity among them is established using the edge list. The *Gephi* tool is used for generating the complete graph structure as described above and the finalized network is shown in Fig. 2. The digraph has n = 4,065, e = 7,909, and 916 bus routes in total.

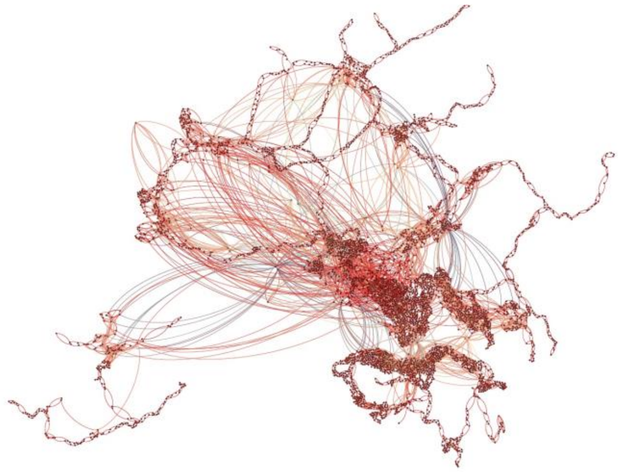

Fig. 2. The structure of the BTN in HK based on the geographical locations (Lat-Long) of bus stops.

2.2. Degree (δ) of Nodes

For a digraph, the in-degree (δ_{in}) is defined as the total number of edge heads injected to a particular node and the out-degree (δ_{out}) is defined as the total number of edge tails connected to the node. Hence, the total degree of a node is $\delta_{total} = \delta_{in} + \delta_{out}$. By measuring the directed in and out degrees of the network, it is found that every node is typically 2-connected in the Hong Kong BTN. Table 1 illustrates the list of nodes with the highest in- and out-degrees. They are regarded as transfer hubs in the BTN.

Table 1. The transfer hubs in HK BTN with the largest in-degree and out-degree.

No	Node Name	δ_{in}	Node Name	δ_{out}
1	Western harbour crossing	23	Lantau link toll plaza	25
2	Lantau link toll plaza	23	Western harbour crossing	17
3	Tate's cairn tunnel	20	Tate's cairn tunnel	15
4	Old wan chai police station	14	Cross harbour tunnel	15
5	Cross harbour tunnel	14	Immigration tower	15
6	Beacon heights	13	Tai lam tunnel	14
7	Cheung on bus terminus	12	Huanggang	12
8	Tuen mun road bus-bus interchange	12	Cheung on bus terminus	11

3. Node Degree Distribution Analysis

Degree distribution corresponds to the probability of finding a node with degree k or the probability distribution of node degrees over the complete network. Barabasi *et al.* [7] showed that if the degree distribution of a network follows power law distribution, then the network is scale-free. This is in contrast to some typical random networks with a binomial or Poisson node degree distribution as proposed by Erdos-Renyi [2]. Mathematically, the power law distribution is defined as

$$P(k) \propto k^{-\alpha} \qquad \text{for } k > k_{\min}, \qquad (1)$$

where $2 < \alpha < 3$ is the Pareto index and k is the node degree.

3.1. Fitting Power Law to Empirical Data

Consider

$$P(k) = C\,k^{-\alpha}, \qquad (2)$$

where C is a constant. Apply log on both sides, we have

$$\log P(k) = -\alpha \log k + \log C\,. \qquad (3)$$

Hence, on log-log scale, the degree distribution of power law follows a straight line with slope $-\alpha$. From (3), it is evident that (2) is valid only for $k > k_{\min}$ and $\alpha > 1$. After normalizing the distribution, we find that $C = (\alpha - 1)/k_{\min}^{-\alpha+1}$. Hence (2) is reduced to

$$P(k) = ((\alpha - 1)/k_{\min})\,(k/k_{\min})^{-\alpha}\ . \qquad (4)$$

Clauset *et al.* showed that the value of $k_{\min}$, which is the lower bound on power law is estimated using the Kolmogorov-Smirnov (KS) test [12], the value of α, which is the scaling parameter is estimated based on maximum-likelihood estimation (MLE). After calculating $k_{\min}$ and α, we perform goodness-of-fit tests between the empirical data distribution and the hypothesized power law distribution and compute the corresponding '*p*-value'. Typically, if the value of $p \geq 0.1$, power law distribution is a plausible hypothesis. By fitting our data to the power law distribution, we found that $\alpha = 3.5$, $k_{\min} = 4$ and the *p*-value < 0.1, which indicates that the network does not follow power law distribution and hence does not behave as a scale-free network. Fig. 3 shows the in- and out-degree distributions of the BTN on a double logarithmic scale.

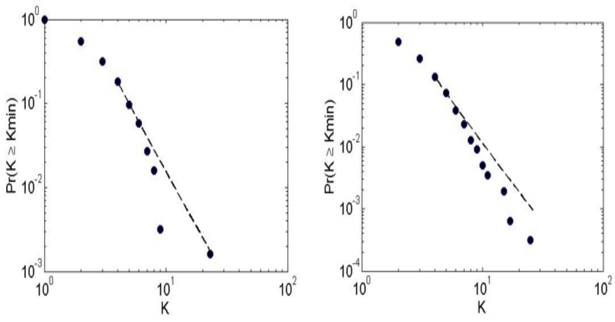

Fig. 3. In-degree and out-degree distributions of the BTN in HK on a log-log scale.

3.2. The Supernode Concept in BTN

As discussed in Section 1, the main idea behind the supernode concept is to consider closely associated bus stops in a BTN as one single node in the graph. The concept of supernode is considered important because it is a more accurate way of looking at the network from a passenger's perspective. Keeping this in mind, a collection of nearby stops is treated as a transfer hub in the PTN. When we look at the geographical positions of bus stops in Fig. 1, it is obvious that sets of bus stops are typically separated by a few meters and are of walkable distance in geographically constrained cities like Hong Kong (Fig. 4 shows several examples). We aim to combine two or more of such nodes together to form a 'Supernode' in the BTN according to some conditions.

Fig.4. Examples of supernode.

<u>Conditions of forming a supernode</u>

As mentioned in Section 2.1, every node in the network is identified by a unique NID, NN and Lat-Long information. Below we discuss the necessary conditions for defining a supernode.

Condition 1: Node_Names and their directions of connectivity

In the Hong Kong BTN, typically if the bus stops have the same name and a different ID, it indicates that the stops are located on either sides of the road, and hence are opposite to each other. Since such nodes are well within walking distance, we combine such nodes together as one supernode.

Condition 2: Geographic distance between nodes

Here we calculate the geographic distance d_{ij} between two nodes and check if $d_{ij} < d_{th}$, where d_{th} represents a certain distance threshold. If d_{ij} is less than d_{th}, then the nodes are combined as a supernode. In our study, the distance threshold is set to be 100 m, which is mostly considered as a walkable distance.

By considering the above conditions, the supernode concept was implemented in the HK BTN, and we verify its scale-free property again after such implementation. For a distance threshold of 100 m, the network structure is re-defined with the new set of in-degree and out-degree, and the corresponding degree distribution is plotted in Figure 5.

Fig. 5 shows the in-degree and out-degree distributions of the HK BTN after the supernode implementation. From the figure, it is more evident that the network structure now closely behaves like a scale-free network with $\alpha = 3.5$, $k_{\min} = 4$, and p-value > 0.1 for KS test, which confirms that the new data set with supernode implementation satisfies the scale-free property. In addition, the network shows an average node degree of 4 compared to 2 before considering the supernodes, i.e., a node is now typically 4-connected rather than 2-connected.

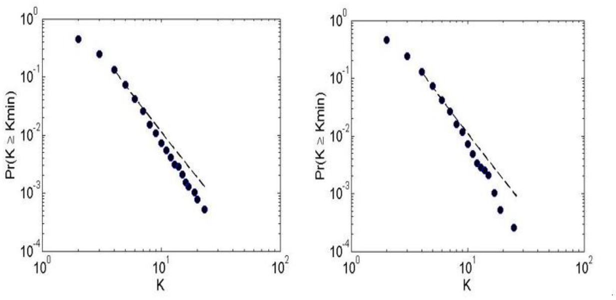

Fig. 5. In-degree and out-degree distributions of the HK BTN on a log-log scale with the supernode implementation.

Table 2 shows the in-degree and out-degree statistics before and after the supernode implementation in the HK BTN. Specifically, we find that the number of high-degree nodes (e.g., from degree 10 to 20) have significantly increased after re-defining the network structure under the supernode concept. As a result, it causes the network to behave like a scale-free network.

Table 2. In- and out-degree statistics before and after implementing the supernode concept.

In-degree	Number of Nodes	Out-degree	Number of Nodes
23	2	25	1
20	1	17	1
14	4	15	4
13	1	14	1
12	2	12	1
11	3	11	6
10	7	10	6

In-degree (Supernode)	Number of Nodes	Out-degree (Supernode)	Number of Nodes
23	2	25	1
20	1	19	1
19	1	17	2
17	1	15	4
16	1	14	2
15	2	13	1
14	3	12	2
13	1	11	6
12	4	10	9
11	5	-	-
10	6	-	-

4. Conclusion

In this paper, we have analyzed the topographical structure of the BTN in Hong Kong and evaluated its scale-free property. With the standardized graph representation, the network initially does not behave as scale-free. However, with the implementation of the proposed supernode concept in the BTN, we found that the network plausibly behaves as a scale-free network. The presence of supernodes in a scale-free bus transport network provides convenient switching points that facilitate efficient routing and hence reduce the average shortest path between any two nodes in the network. It is therefore of practical relevance to study the impact of network structure on traffic performance and to identify key factors and parameters that affect performance of public transportation systems.

Acknowledgement

This work is partially supported by the Early Career Scheme (Project No. 25200714) established under the University Grant Committee of the Hong Kong Special Administrative Region, China; the National Natural Science Foundation of China (Project No. 61401384); and The Hong Kong Polytechnic University (Projects 4-ZZCZ, G-YBK6, and G-YN17).

References

[1] Passenger Journeys by Fixed Route Public Transport Operator, Transport Department, HKSAR Government, 2016.

[2] P. Erdos, A. Renyi, "On Random Graphs," *Publ. Math.* (Debrecen) 6, pp. 290–297, 1959.

[3] R. Albert and A. -L. Barabasi, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, pp. 47–94, 2002.

[4] A. Chatterjee, M. Manohar and G. Ramadurai, "Statistical Analysis of Bus Networks in India," *arXiv preprint*, arXiv: 1509.04554v2, pp. 1–18, 2015.

[5] H. Zhang, P. Zhao, J. Gao and X. Yao, "The Analysis of the Properties of Bus Network Topology in Beijing on Complex Networks," *Mathematical problems in Engineering*, vol. 2013, pp. 1–6, 2013.

[6] H. Soh, S. Lim, T. Zhang, X. Fu, G. Lee, T. Hung, P. Di, S. Prakasam, L. Wong, "Weighted complex network analysis of travel routes on the Singapore public transportation system," *Elsevier Physica*, vol. 389, pp. 5852 –5863, 2010.

[7] C. Von Feber, T. Holovatch, Y. Holovatch and V. Palchykov, "Public Transport Networks: empirical analysis and modelling," *Eur. Phys. J. B 68*, pp. 261–275, 2009.

[8] J. Sienkiewicz and J. A. Holyst, "Statitical Analysis of 22 Public Transport Networks in Poland,", *arXiv preprint*, arXiv: physics/0506074v2, pp. 1–11, 2005.

[9] [Online] https://data.gov.hk/en/

[10] A.-L. Barabasi and R. Albert, "Emergence of scaling in random networks," *Science*, 286(5439):509–512, 1999.

[11] *User Handbook on Datun Transformations Involving WGS84*, International Hydrographic Organization, 3rd edition, No. 60, 2003.

[12] A. Clauset, C. R. Shalizi and M. E. J. Newman, "Power-Law distribution in Empirical Data," *arXiv preprint*, arXiv: 0706.1062v2, pp. 1–22, 2009.