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

A Network Analysis of World's Metro Systems

Xingtang Wu†, Chi K. Tse‡, Hairong Dong†, Ivan W. H. Ho‡ and Francis C. M. Lau‡

†School of Electronic and Information Engineering, Beijing Jiaotong University, China
‡Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong, China
Email: 14111051@bjtu.edu.cn, cktse@ieee.org

Abstract—Metro systems carry a large volume of commuters around in major cities and their efficiency depends on multiple factors including the connectivity of subway lines, train schedules, passenger distributions, management quality, etc. In this paper, we study the network properties of five metro systems (Beijing, Hong Kong, London, Paris and Tokyo) and compare their relative network efficiency in terms of two proposed metrics, namely, average station density and station load. It is shown that among the five systems, the Tokyo system has the shortest characteristic path length (shortest average travel distance between two stations), as well as highest efficiency in carrying passengers around the city. Furthermore, the London metro has a better tolerance to faults in a local scale, and the Paris system outperforms others in terms of level of convenience to commuters due to its high station density and low load.

1. Introduction

Rapid transit systems, often called metro or subway systems, are transportation systems carrying the largest volume of commuters in major cities, and their reliability, efficiency, safety, levels of comfort, convenience and accessibility are often perceived by travellers and local commuters as indicators of the quality of public transportation of the cities [1]. Major cities, due to increasing traffic demands and ever-extending city coverage, are continuously expanding their metro networks, resulting in complex subway systems that possess high station densities and intricate inter-station couplings [2]. Design and scheduling of metro systems to optimize performance has become important considerations in the development of public transportation systems. Moreover, the study of networks, under the notion of *complex networks*, has recently become popular due to the intriguing discovery of a number of universal properties in various physical and man-made networks [3, 4] and the promising applications that have been developed in various practical fields such as communications, power systems, finance, disease control, etc. [5]–[10]. Results from complex networks research are highly relevant to the study of transportation, especially in the provision of appropriate analytical tools for characterizing the structure of metro systems which are practical forms of networks and for understanding the operations of a complex system such as metro systems [11, 12]. Furthermore, the huge investment in this transportation infrastructure and the impact to the public certainly justify a more thorough investigation of the factors affecting performance, thus allowing a more informed planning and design for future development.

The cross-disciplinary study of subway systems from a perspective of complex networks is still relatively rare. The earliest work reported by Latora and Marchiori [13] showed that the Boston subway network exhibited the small-world property and introduced the concept of network efficiency to give useful insights on the general characteristics of real transportation networks. In the work of Derrible and Kennedy [14], most metros were found to exhibit scale-free and small-world structure. Also, Angeloudis and Fisk [2] studied 20 subway networks using a 'toy' model and showed that these networks, with high connectivity and low maximum vertex degrees, provide robustness to random attacks. In the work of Lee *et al.* [15], the statistical properties of the Metropolitan Seoul subway network were analyzed, taking the passenger flow as the weight of the edge and arriving at a power-law weight distribution. Furthermore, Yang *et al.* [5] combined node degree and betweenness to assess the node importance, and showed that a scale-free transit network exhibited a relatively high fault tolerance to random failure but a relatively low degree of connection reliability against malicious attack.

In this paper, five subway networks are studied via analyzing some network parameters such as degree distribution and network efficiency, the aim being to identify the factors affecting performance.

2. Topological properties of subway network

A complex network with N nodes can be represented as a graph $G = (N_d, l)$, where $N_d = \{n_1, n_2, ..., n_N\}$ denotes the set of nodes, and $l = \{l_1, l_2, ..., l_k\}$ denotes the set of links. A graph G can be fully described by an adjacency matrix A, which is an $N \times N$ matrix whose entry $a_{ij}(i, j = 1, ..., N)$ equals to 1 if there exists a link between nodes i and j, and zero otherwise. In this paper, a node is a subway station, if two stations are directly connected by a track, then they are connected by a link.

2.1. Characteristic path length

Shortest path length, denoted as d_{ij}, is the shortest length from nodes i to j, which plays an important role in transportation and communication networks. Suppose one needs to commute from one station to another by subway: the shortest path provides an optimal pathway in the sense that one would achieve a fast transfer, saving time and resources. A measure of the typical separation between two nodes in a complex network is given by the characteristic path length, also known as *average path length*, which is defined as the mean shortest path lengths over all pairs of nodes [3]:

$$L = \frac{1}{N(N-1)} \sum_{i,j \in N, i \neq j} d_{ij} \tag{1}$$

This parameter directly indicates the global connectivity of a network. A smaller value of L represents smaller topological distance between any two nodes and better connectivity of the whole network.

2.2. Clustering coefficient

Clustering coefficient C, also known as transitivity, is a typical property of acquaintance networks, where two individuals with a common friend may know each other. One definition of C, introduced by Watts and Strogatz [3], is given as follows. A quantity c_i (local clustering coefficient of node i) is first defined to describe how likely a_{jm}=1 for two neighbors j and m of node i. It is defined as the ratio between e_i and $\frac{k_i(k_i-1)}{2}$, in which e_i denotes the actual number of edges between the neighbors of node i, i.e.,

$$c_i = \frac{2e_i}{k_i(k_i-1)} = \frac{\sum_{j,m} a_{ij} a_{jm} a_{mi}}{k_i(k_i-1)} \tag{2}$$

The clustering coefficient of a graph is the average of c_i over all nodes:

$$C = \frac{1}{N} \sum_{i \in N} c_i \tag{3}$$

Thus, $0 \leq c_i \leq 1$. The clustering coefficient indicates the local clustering property and shows the fault tolerance characteristic. Taking the subway network as an example, when one track is out of function, the traffic will not be affected if the neighboring stations are connected. Thus, a larger value of C denotes a better tolerance to fault in a local scale.

2.3. Network (Structural) Efficiency

Efficiency E, introduced by Latora and Marchiori [13], is a measure of how efficient information is exchanged over the network. Denoted as ϵ_{ij}, efficiency of transfer from nodes i to j is taken as being inversely proportional to the shortest path length, i.e., $\epsilon_{ij} = \frac{1}{d_{ij}} \forall i, j$, and the network efficiency E is defined as:

$$E(G) = \frac{1}{N(N-1)} \sum_{i \neq j} \epsilon_{ij} = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \tag{4}$$

Note that $E(G)$ is the global efficiency of the whole network and is denoted as E_{glob}. Also, $E(.)$ can be defined to characterize the local properties of G by evaluating the efficiency of G_i, the subgraph consisting of the neighbors of node i but excluding node i. The local efficiency E_{loc} is then defined as the average efficiency of all subgraphs:

$$E_{\text{loc}} = \frac{1}{N} \sum_{i \subset G} E(G_i) \tag{5}$$

E_{loc} plays a similar role as C, and tells how efficient the communication between the neighbors of i is in the absence of node i, reflecting the robustness of local connection when node i is removed.

However, this definition of E is not fully consistent with the subway operation. In subway networks, segments of some lines overlap, thus affecting transportation efficiency. To correct this, if nodes i and its neighbor j are connected by multiple edges, we scale the link connecting the two nodes by a factor w_{ij} and use the scaled link to compute d_{ij}, i.e.,

$$w_{ij} = \frac{1}{n} \tag{6}$$

where n is the number of edges between station i and its neighbor station j. Then, E is calculated based on the weighted network structure.

2.4. Average station coverage area and load

In order to evaluate the average distance from a random passenger to a subway station and the average passenger load of a station, we propose two parameters, namely, *average station coverage area* (ASCA) and *average station load* (ASL). Here, we define ASCA as as the ratio of subway network area S_{all} and the number of station,i.e., ASCA = $\frac{S_{all}}{N}$. Thus, ASCA reveals the average area served by a station, or equivalently, the average distance to a subway station for passengers, and the station density. Moreover, ASL is defined as the ratio of average daily passenger flow P and the number of stations, i.e., ASL = P/N, reflecting on the average crowdedness of the stations.

3. Statistical results

In this paper, the subway networks in Beijing, Hong Kong, London, Paris and Tokyo are studied. Basic information of these subway networks are listed in Table 1.

Table 1: Basic data of subway scale (as of 2015)

City	Number of stations	Number of lines
Beijing	274	18
Hong Kong	85	10
London	356	13
Paris	295	15
Tokyo	205	13

Figure 1 shows the shortest path length distribution of the five subways. It is observed that they basically follow

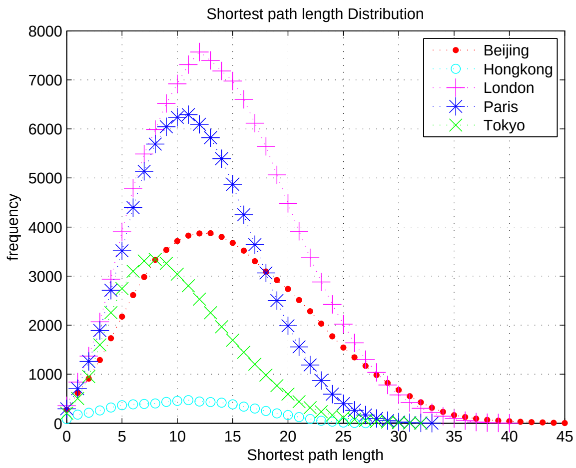

Figure 1: Shortest path length distribution

Table 2: Characteristic path lengths of subway networks

City	Beijing	HK	London	Paris	Tokyo
N	274	85	356	295	205
L	15.02	10.97	14.03	11.77	**10.13**

the Γ distribution and Hong Kong has the smallest value of network diameter, which is defined as the maximal shortest path length of a network.

The characteristic path lengths L for the five systems are listed in Table 2, which shows that Tokyo offer the shortest characteristic path length, and Hong Kong and Paris having slightly longer characteristic path lengths.

Clustering coefficients C are calculated and listed in Table 3, from which we can see that the London subway has a relatively bigger C, implying a better tolerance to faults in a local scale.

Efficiency E, based on weighted edges as explained in Section 2.3, is compared in Table 4. It can be shown from the result that all those five subway networks behave less efficient in the topological level compared to a fully connected network (which has a theoretical efficiency of 1). This is because the number of edges $Q \ll N(N-1)/2$ and the neighbors of most nodes are isolated from each other. From the values of E we can see that the Tokyo and Hong Kong subways perform better than others in the global scale while the Tokyo and London systems perform better in the local scale.

In this paper, the area served by a subway network is conveniently taken as a rectangle, whose edges are defined by the position of the farthest stations in the four directions. For example, for the Beijing subway shown in Fig. 2, the rectangular boundaries are decided by the farthest stations: Nanshao, Tiangongyuan, Suzhuang and Lucheng. For a fair comparison, the sea areas within the areas covered by the subway are removed for Tokyo and Hong Kong. Further, as Hong Kong is a mountainous city, where only 25% of the defined rectangle is inhabited, we adjust the effective area served by the subway accordingly. Table 5 lists the areas and passengers served by the individual subway systems.

In order to see the station density variation, we divide the area served into five concentric rectangular regions (rectan-

Table 3: Clustering coefficients of subway networks

City	Beijing	HK	London	Paris	Tokyo
C	0.0024	0.0059	**0.0409**	0.0163	0.0285

Table 4: Efficiency of subway networks

City	Beijing	HK	London	Paris	Tokyo
E_{glob}	0.1012	0.1526	0.1261	0.1143	**0.1560**
E_{loc}	0.0024	0.0058	**0.0339**	0.0146	0.0319

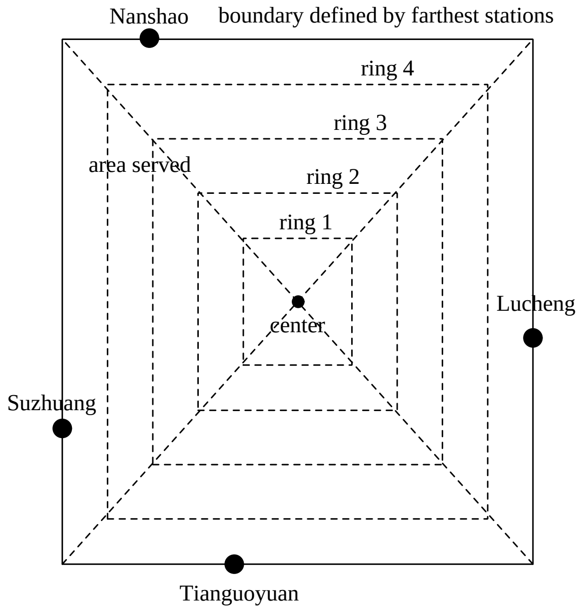

Figure 2: Subway map representation. Boundary defined by 4 farthest stations in 4 directions. Station names refer to Beijing system.

Table 5: Data on areas and passengers served by subways

City	Beijing	HK	London	Paris	Tokyo
Area served (sq. km)	3217.0	177.2	1920.0	347.0	594.3
Passengers per day (million)	10.876	4.490	8.245	4.130	8.500

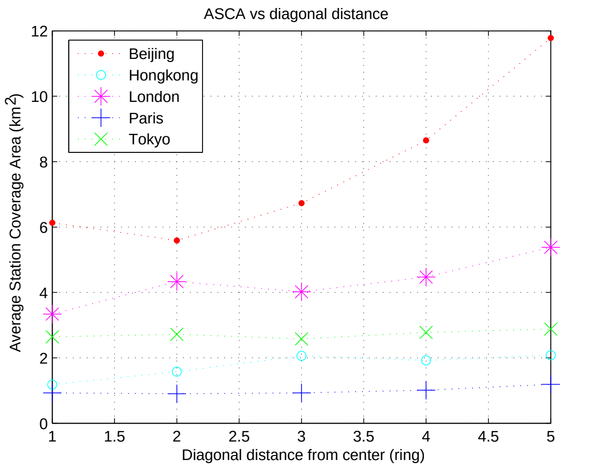

Figure 3: ASCA vs. diagonal distance from center (ring). x-axis is ring number, 1 being the first inner ring and 5 being the entire area.

Table 6: Average station load

City	Beijing	HK	London	Paris	Tokyo
ASL /day (x1000)	39.69	52.82	23.16	**14.00**	41.46

gular rings), along the diagonal direction. Fig. 3 shows the ASCA versus the diagonal distance from center (ring). For instance, ASCA at ring = n is the ASCA of the inner area within the nth rectangular ring. Also the values of ASL are shown in Table 6. We see that the Paris subway has relatively small ASCA and ASL, and therefore has higher density and is more convenient for passengers. In addition, stations in Hong Kong, Paris and Tokyo are basically distributed uniformly over the city.

4. Conclusion

The topological structure of five subway networks are studied in terms of the characteristic path length, clustering coefficient, and network efficiency. We propose two parameters, namely, average station coverage area (ASCA) and average station load (ASL), to evaluate the station density and the level of convenience to passengers. Among the five subway networks, the Hong Kong subway has the smallest characteristic distance, and the Tokyo subway has the highest topological efficiency. The London subway has a larger value of clustering and local efficiency, suggesting that it has a better tolerance to fault in a local scale. The Paris subway offers the highest level of convenience to passengers due to the low ASCA and ASL.

Acknowledgement
This work is supported by National Natural Science Foundation of China under Grant 61322307.

References

[1] S. Su, T. Tang and C. Roberts, "A cooperative train control model for energy saving," *IEEE Trans. Intelligent Transportation Syst.*, vol. 16, no. 2, pp. 622–631, Apr. 2015.

[2] P. Angeloudis and D. Fisk,"Large subway systems as complex networks," *Physica A*, vol. 367, pp. 553–558, July 2006.

[3] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, pp. 440–442, June 1998.

[4] A. L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, Oct. 1999.

[5] J. Yang, C. Yao, and W. Ma, "A study of the spreading scheme for viral marketing based on a complex network model," *Physica A*, vol. 389, no. 4, pp. 859–870, Feb. 2010.

[6] M. Small, D. M. Walker, and C. K. Tse, "Scale free distribution of avian influenza outbreaks," *Physical Review Letters*, vol. 99, 188702, Nov. 2007.

[7] C. K. Tse, J. Liu, and F. C. M. Lau, "A network perspective of the stock market," *J. Empirical Finance*, vol. 17, no. 4, pp. 659–667, Sept. 2010.

[8] J. Wu, C. K. Tse, F. C. M. Lau and I. W. H. Ho, "Analysis of communication network performance from a complex network perspective," *IEEE Trans. Circ. Syst. I: Reg. Papers*, vol. 60, no. 12, pp. 3303–3316, Dec. 2013.

[9] D. He, R. Lui, L. Wang, C. K. Tse, L. Yang, and L. Stone, "Global spatio-temporal patterns of influenza in post-pandemic era," *Scientific Reports*, vol. 5, article 11013, DOI: 10.1038/srep11013, June 2015.

[10] X. Zhang and C. K. Tse, "Assessment of robustness of power systems from a network perspective," *IEEE J. Emerg. Select. Topics Circ. Syst.*, vol. 5, no. 3, pp. 456–464, Sept. 2015.

[11] Y. Bar-Yam, S. R. McKay, and W. Christian, *Dynamics of Complex Systems (Studies in Nonlinearity)*, Computers in Physics, 1998.

[12] A. L. Barabási, "The network takeover," *Nature Physics*, vol. 8, pp. 14–16, 2012.

[13] V. Latora and M. Marchiori, "Is the Boston subway a small-world network?" *Physica A*, vol. 314, pp. 109–113, 2002.

[14] S. Derrible and C. Kennedy, "The complexity and robustness of metro networks," *Physica A*, vol. 389, no. 17, pp. 3678–3691, Sept. 2010.

[15] K. Lee, W. S. Jung, J. S. Park, and M. Y. Choy, "Statistical analysis of the Metropolitan Seoul subway system: network structure and passenger flows," *Physica A*, vol. 387, no. 24, pp. 6231–6234, 2008.