Understanding the Structure of Urban Bus Networks: The *C*-Space Representation Approach

Lijun Sun[1]; Yang Lu[2]; and Der-Horng Lee[3]

[1]Future Cities Laboratory, Singapore-ETH Centre, 1 Create Way, #06-01 CREATE Tower, Singapore 138602. E-mail: lijun.sun@ivt.baug.ethz.ch
[2]Singapore-MIT Alliance for Research and Technology (SMART), 1 Create Way, #09-01 CREATE Tower, Singapore 138602. E-mail: luyang@smart.mit.edu
[3]Department of Civil & Environmental Engineering, National University of Singapore, 1 Engineering Drive 2, E1A #07-03, Singapore 117576.
E-mail: dhl@nus.edu.sg

Abstract

The application of network science has greatly improved our understanding about the structure and properties of public transport network (PTN). Using tools and methods provided from network science, several standardized representations have been developed and analyzed to understand the topological structure of PTN from different aspects. In this paper, we present an empirical study on the bus network in Singapore. Different from previous studies based on a direct mapping of PTN on road networks, we focus on the weighted *C*-space representation of public transport networks to better characterize the interaction of different bus services. By exploring the statistical properties of this representation, we find that the bus network in Singapore exhibits similar patterns as many other cities in terms of distributions of edge weight, vertex degree and vertex strength. However, a superlinear scaling is found when we examine the dependency of a vertex's strength on degree, indicating that PTN in Singapore provides people with more alternatives to travel and transfer. The *C*-space representation also enables us to further explore the interactions between services by examining the community structure and the corresponding reverse mapping on road networks. This helps us to identify those crucial 'bridges' in a city, which is shared by different communities and connect diverse regions.

INTRODUCTION

Public transport networks (PTN) are one the most important infrastructure layers in urban functional systems, providing people with affordable access to education, employment, markets and other key urban services with less environmental impact. A comprehensive understanding of the structure and properties of PTN will benefit urban planning, traffic operation and infectious disease control. In the last decade, there is an increasing number of studies on the topological properties of transportation networks using tools and methods provided from complex network theory (Barthélemy 2011; Lin and Ban 2013). For example, researchers have investigated the structure of urban street networks using the primal structure and the corresponding dual structure (Porta et al. 2006a; Porta et al. 2006b).

A body of literature has investigated the structural properties of PTN in diverse cities − such as Beijing, Shanghai, Berlin, Paris, Singapore − using the emerging network theory. In order to capture the different and unique topological structure and properties, previous studies have investigated different standardized network representations of PTN, such as the *L*-space and *P*-space representations (Sienkiewicz and Hołyst 2005; Von Ferber et al. 2009). In general, *L*-space shows the simplified graph of PTN. Vertices in the *L*-space network represent bus stops and an edge is created by connecting those consecutive bus stops if there is at least one bus route. The *P*-space is also constructed by bus stops as vertices; however, for each bus route, all the stops are fully connected by unweighted edges. Using bus network data in the Chinese city Shijiazhuang from 1996 to 2008, Yan and Wang (2009) studied the topological properties of *P*-space representation in temporal scale. Taking advantage of the temporal evolution of the PTN in Shijiazhuang, the authors found that PTN in the *P*-space is a small-world and accelerated growing network. Different from *L* and *P*-space, the *C*-space network is a complementary projection, in which vertices denote bus services and an edge exists if two services have at least one overlapping stop. Xu et al. (2007) conducted the first comprehensive investigation on the *C*-space structure of PTN. The authors explored and compared the structure properties of *C*-space networks of PTN from Beijing, Shanghai and Nanjing. The authors stated that the weighted *C*-space network can better characterize the structure of PTN from the information perspective. Using the real travel data in Singapore's PTN, Soh et al. (2010) created a new representation with bus stops as vertices; however, the directed edges are defined by number passengers travel from one stop to another. As a result, the authors created a directed and weighted network using real origin-destination passenger flow.

In this paper, we analyze the structural properties of the *C*-space representation of the bus network in Singapore. We show some unique characteristic of this empirical PTN, which can only be interpreted from the *C*-space representation. The contributions of this paper are twofold: first, we explored the statistical properties of the weighted *C*-space representation and compared with results with other cities. We found that PTN in Singapore has a higher degree of overlapping to ensure enough connectivity/accessibility; second, we show that the *C*-space representation provides a community structure with overlapping at meaningful spatial scale, which cannot be obtained from a representation with bus stops as vertices. The reminder of this paper is organized as follows: in section 2, we introduce the original PTN (bus) in Singapore and a simplified version by aggregating two-way services and bus stops correspondingly. Then we construct the dual network based on the simplified PTN; in section 3, we explore the statistical properties of the dual network by looking at edge weight, degree, strength, and community structure; and finally section 4 summarizes and concludes this study and provides the outlook for future works.

CASE STUDY: BUS NETWORK IN SINGAPORE

Bus Network

The case study is based on bus network of Singapore in 2014. The detailed information of bus systems is obtained from StreetDirectory (http://www.streetdirectory.com). Figure 1 shows the spatial network structure, in which the red dots represent bus stops and the gray lines are links in the corresponding road network. The thickness of a link captures the number of bus routes passing through it. In total, this network is built on 428 bus routes, consisting of 4620 bus stops and 6291 links. The longest bus service is route SBS-63, containing 99 bus stops, while the shortest bus services have only 4 stops. The average number of bus stops per route is 39. Before analyzing the network, we first simplify it by aggregating two opposite routes of the same service. This is done by keeping only one of the two opposite routes. In so doing, we combine pairs of bus stops that are opposite to each other on the road.

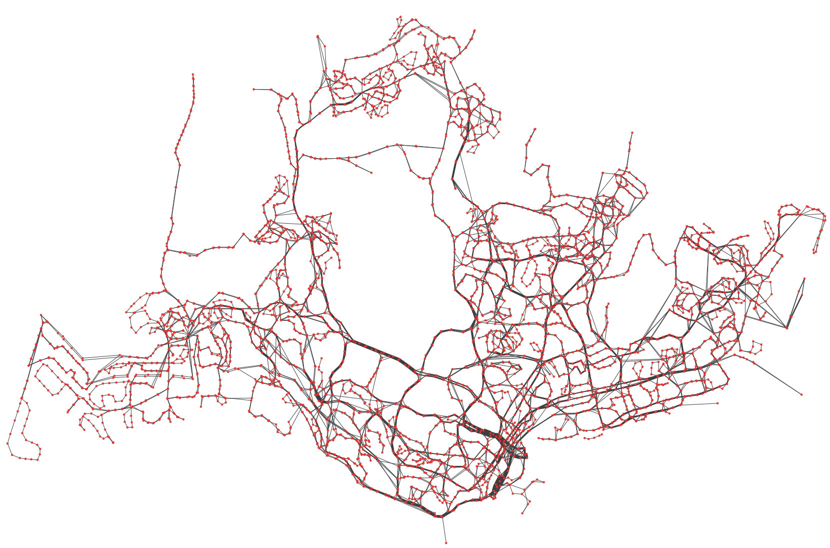

Figure 1. PTN network (bus) in Singapore

The simplified PTN network consists of 2611 vertices (combined bus stops) from 309 bus services. Since some routes are loop services operating on one direction only, the number of services is higher than 428/2=214. We next study the properties of this network based on the complementary projection to service vertices, which is referred as the *C*-space network (Von Ferber et al. 2009).

The *C*-space Representation

We construct the dual network based on the simplified PTN. Different from the other representations in the *L*-space and *P*-space, a vertex i in the *C*-space network denotes a bus service. Therefore, as mentioned, *C*-space network provides a complementary projection of a PTN based on services instead of bus stops. The *C*-space representation can better capture the connectivity/accessibility of a PTN in which one bus stop serves multiple services. This representation is particularly useful to measure the number of transfers between different services in a PTN. A comprehensive illustration of different standardized representations in *L*-space, *P*-space, *B*-space and *C*-space can be found in Von Ferber et al. (2009). We donate V_i as the set of bus stops along service i in the PTN. Similar to previous study (Xu et al. 2007), in this representation we create an edge between service i and service j if $V_i \cap V_j \neq \varnothing$. Therefore, in the *C*-space representation, the existence of an edge (i, j) indicates that there is overlap between those two services in the PTN. Finally, we produce an undirected and weighted network, which captures the dual structure of PTN to study its properties. The dual network contains 309 vertices and 15,152 edges.

STATISTICAL PROPERTIES OF THE *C*-SPACE REPRESENTATION

Edge Weight

As defined, in the dual network the edge linking two vertices represents the overlap of two bus services. To quantify the degree of overlapping of two bus services, we define edge weight of edge (i, j) as the number of overlapping/common bus stops between service i and j, i.e., $w_{ij} = |V_i \cap V_j|$. Therefore, w_{ij} measures the number of overlapping stops of two bus routes. The average edge weight is $\langle w_{ij} \rangle = 3.8$. The minimal edge weight in the *C*-space network is one, characterized by the paired services sharing only one common stop (e.g., two services sharing only the same departure depot or terminus). The maximal weight observed in the network is 42. To explore the statistical property of edge weight, we first measured its cumulative distribution $P(w)$. In general, a cumulative distribution $P(x)$ is defined as the probability of observing cases with attribute $z \geq x$ (e.g., degree and strength):

$$P(x) = \sum_{z \geq x} p(z). \tag{1}$$

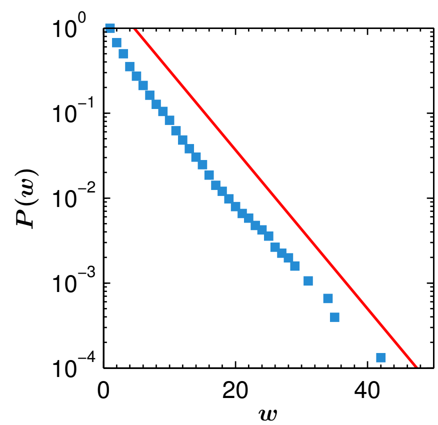

Figure 2. Cumulative edge weight distribution $P(w)$

Figure 2 shows the cumulative edge weight distribution $P(w)$. As can be seen, $P(w)$ is well-characterized by an exponential decaying distribution $P(w) \sim \exp(-\beta w)$ with $\beta = 0.215 \pm 0.003$ (the red line shows an exponential curve with $\beta = 0.215$ as a guide), indicating that most paired services share limited number of common bus stops. Similar edge weight distributions are observed in other *C*-space representations of PTN in Beijing, Shanghai and Nanjing (Xu et al. 2007). The bounded nature of edge weight w also suggests that

Degree and Strength

In complex networks, the strength of a vertex is defined as the sum of weights of its adjacent edges:

$$s_i = \sum_{j \in N(i)} w_{ij} \ , \tag{2}$$

where $N(i)$ is the set of neighbors of vertex i. Therefore, s_i captures the overall degree of overlapping that service i connects with other services in the PTN.

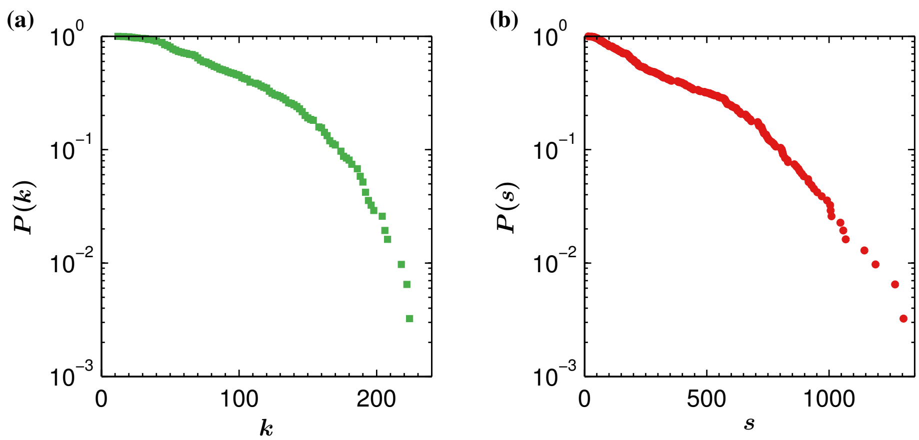

Figure 3. Cumulative distributions of degree $P(k)$ and strength $P(s)$

Figure 3 shows the cumulative distributions of degree and strength in the *C*-space representation in log-linear plots.

We find that $P(k)$ in *C*-space network exhibits an exponential decaying tail which is observed in the PTN in Berlin, London and Los Angeles (Von Ferber et al. 2009). Moreover, $P(s)$ also displays an exponentially decaying tail, indicating that the degree of overlapping with other bus services is bounded. This observation is also in accordance with the design principle of a PTN: one should make bus services more diversely distributed and avoid too many parallel services. We next study the dependence of s on k, which is a crucial characteristic of weighted networks (Barrat et al. 2004).

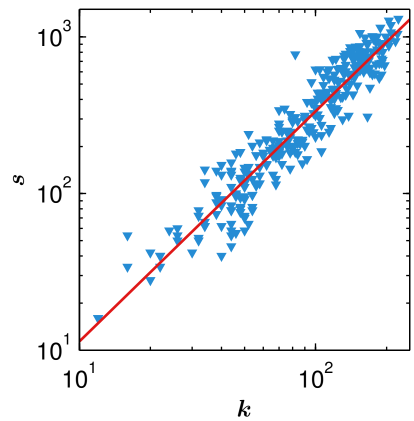

Figure 4. Scatter plot of vertex degree and strength in the *C*-space network

Figure 4 shows the scatter plot of vertex degree k and vertex strength s of the C-space representation of PTN in Singapore. Different from the linear relation $s \sim k$ observed in Beijing, Shanghai and Nanjing (Xu et al. 2007), the dependency in Singapore's network is well characterized by a superlinear scaling:

$$s \sim k^{\beta} \text{ with } \beta > 1, \quad (3)$$

where $\beta = 1.47 \pm 0.03$ ($R^2 = 0.90$), indicating an increasing degree of overlapping with number of connected services in Singapore. In other words, the more connected one service is, the more overlapping stops those connected services have on average ($s/k \sim k^{\beta-1} > 1$). Compared to a randomized weighted network with $\beta = 1$, this finding also demonstrates a collective growth effect --- a more highly connected bus service also has more common stops with other services. This structural property provides people with more options with less waiting time to travel to a destination and to transfer from one bus service to another. As a result, the PTN in Singapore has higher resilience against attacks and higher tolerance of interruptions, disruptions and failures. However, such tolerance comes with a high cost by maintaining services with overlapping functions, leading to a waste of resources when supply is limited.

Table 1 summarizes the descriptive statistics of the C-space representation of PTN in Singapore.

Table 1. Descriptive statistics of the *C*-space network

	Mean $\langle x \rangle$	**Std** $\sigma(x)$	**Min** $x_{\min}$	**Max** $x_{\max}$
Edge weight w	3.8	3.8	1	42
Degree k	98.1	50.7	12	224
Strength s	371.9	289.3	16	1304

Community Structure

The weighted *C*-space representation characterized not only the topological connectivity of bus services in the PTN, but also the degree of similarity by measuring number of overlapping stops. To further understand the property of PTN, we conduct community detection on the corresponding *C*-space network. The community structure of a network is defined as vertex partitions with more connections within themselves than between each other. Essentially, such a community structure can be obtained by maximizing modularity Q (Newman 2004):

$$Q = \frac{1}{2W} \sum_{i,j} \left(w_{ij} - w'_{ij} \right) \delta\left(c_i, c_j \right), \tag{4}$$

where $W = \sum_i s_i$ is the total weight of all edges, w'_{ij} is expected weight estimated from a null model, and $\delta\left(c_i, c_j\right)$ is an indicator function with $\delta\left(c_i, c_j\right) = 1$ if zone i and zone j belong to the same partition and $\delta\left(c_i, c_j\right) = 0$ otherwise. In this study, we use the random null model with $w'_{ij} = s_i s_j / 2W$. To find the pattern of partitions with the highest modularity, we employed the well-established Louvain method (Blondel et al. 2008) on the *C*-space network. Finally, we obtained the best community structure with modularity $Q = 0.34$.

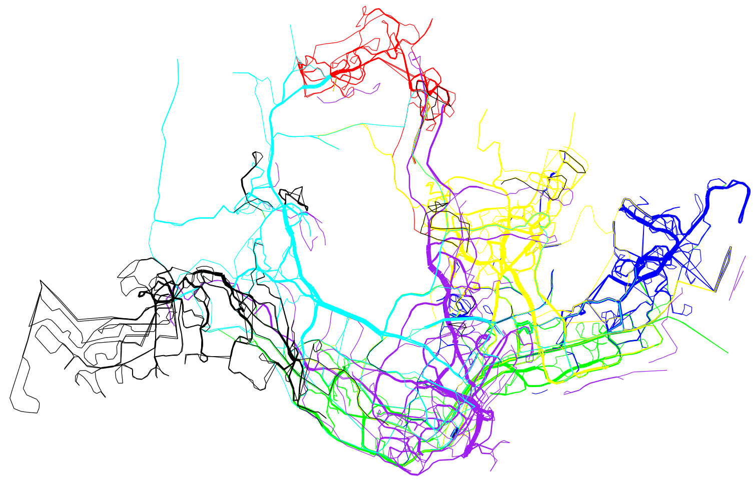

Figure 5. Mapping *C*-space community structure to the original PTN

To see how bus services interact with each other, in Figure 5 we mapped the community structure in *C*-space to the original PTN (see Figure 1). The thickness of each colored link corresponds to the number of services going through the link in community of the same color. In total, seven communities of bus services with more than one service are identified from (shown in different colors). We found that all communities are consistently identified based on their spatial/geographical coverage. As can be seen, most links in the road network serves only one service community (marked with one color); however, there are still some links occupied by different communities with diverse intensity, in particular in the city center. These shared links establish better and more convenient connections to other from one community to another, serving as structural 'bridges' in the original PTN.

The advantage of using *C*-space representation in community detection is that vertices are denoted by bus services themselves instead of bus stops. Therefore, the partitions can better capture the interaction between different but services, while the partitions in *L*-space and *P*-space will only provide clusters for bus stops. However, in reality, many trunk bus services are operated across diverse transportation zones. The resulted community structure from *L* and *P*-space may lack of physical and spatial meaning. On the other hand, applying community detection on stop-based networks prevents us from identifying the overlapping properties services clusters. For example, two services passing through the same road segments may belong to different communities. In other words, the links on road network could serve different services clusters. As a result, links and regions covering multiple communities behave as structural bridges in city and it is crucial to identifying these links and regions for network planning and transit operation.

CONCLUSION

In this paper, we analyzed the structure and properties of the *C*-space representation of Singapore's bus network. In order to capture the interaction intensity among different bus services, we created a weighted *C*-space representation which denotes bus services as vertices and uses the number of common bus stops as edge weight between a pair of services.

By examining the statistical properties of the *C*-space network, we found that the PTN in Singapore are comparable to many other cities globally, such as Beijing, Shanghai, Berlin and Paris, displaying similar patterns in terms of distributions of edge weight, vertex degree and vertex strength. The results indicate that transit networks might be governed by the same design principles. However, when measuring the dependency of strength on degree of all services, the *C*-space network in Singapore exhibits a superlinear scaling instead of linear relations observed in Beijing, Shanghai and Nanjing. Essentially, one of the most important principles of designing a PTN is to ensure enough overlap to maintain certain degree of connectivity/accessibility. A superlinear scaling indicates that PTN in Singapore provides an increasing capacity and tolerance for interruptions, disruptions and other network failures. However, on the other hand, this may cause a waste of recourses (buses) when transit supply is limited.

Taking advantage of the *C*-space structure in which vertices are defined by services, we also studied the community structure to better characterize the interactions with network topology. This approach enables us to identify critical links that serve as structural 'bridges' in the original PTN. The results and methodologies presented in this paper can help us understand the structure of urban transit system, facilitating better design and operation of PTN.

ACKNOWLEDGEMENT

This research was supported by National Research Foundation (NRF) of Singapore, which is the funding authority of Future Cities Laboratory, Singapore-ETH Centre and Singapore-MIT Alliance for Research and Technology.

REFERENCES

Barrat, A., Barthélemy, M., Pastor-Satorras, R., and Vespignani, A. (2004). "The architecture of complex weighted networks." *Proceedings of the National Academy of Sciences of the United States of America*, 101(11), 3747-3752.

Barthélemy, M. (2011). "Spatial networks." *Physics Reports*, 499(1), 1-101.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). "Fast unfolding of communities in large networks." *Journal of Statistical Mechanics: Theory and Experiment*(10), P10008.

Lin, J., and Ban, Y. (2013). "Complex Network Topology of Transportation Systems." *Transport Reviews*, 33(6), 658-685.

Newman, M. E. J. (2004). "Analysis of weighted networks." *Physical Review E*, 70(5), 056131.

Porta, S., Crucitti, P., and Latora, V. (2006a). "The network analysis of urban streets: a dual approach." *Physica A: Statistical Mechanics and its Applications*, 369(2), 853-866.

Porta, S., Crucitti, P., and Latora, V. (2006b). "The network analysis of urban streets: a primal approach." *Environment and Planning B: Planning and Design*, 33, 705-725.

Sienkiewicz, J., and Hołyst, J. A. (2005). "Statistical analysis of 22 public transport networks in Poland." *Physical Review E*, 72(4), 046127.

Soh, H., Lim, S., Zhang, T., Fu, X., Lee, G. K. K., Hung, T. G. G., Di, P., Prakasam, S., and Wong, L. (2010). "Weighted complex network analysis of travel routes on the Singapore public transportation system." *Physica A: Statistical Mechanics and its Applications*, 389(24), 5852-5863.

Von Ferber, C., Holovatch, T., Holovatch, Y., and Palchykov, V. (2009). "Public transport networks: empirical analysis and modeling." *The European Physical Journal B-Condensed Matter and Complex Systems*, 68(2), 261-275.

Xu, X., Hu, J., Liu, F., and Liu, L. (2007). "Scaling and correlations in three bus-transport networks of China." *Physica A: Statistical Mechanics and its Applications*, 374(1), 441-448.

Yan, X., and Wang, M. "Topological properties of public transport networks: a temporal perspective." *Proc., ICCTP 2009: Critical Issues in Transportation Systems Planning, Development, and Management*, ASCE, 1084-1089.