Graph Fourier Transform Centrality for Taipei Metro System

Chien-Cheng Tseng
Department of Computer and Communication Engineering
National Kaohsiung University of Science and Technology
Kaohsiung, Taiwan
tcc@nkust.edu.tw

Su-Ling Lee
Department of Computer Science and Information Engineering
Chang-Jung Christian University
Tainan, Taiwan
lilee@mail.cjcu.edu.tw

***Abstract*—In this paper, the graph Fourier transform (GFT) centrality is used as node importance measure to identify central stations of the Taipei metro system. First, the computation algorithm of the GFT centrality is presented. Then, GFT centrality is employed to analyze Taipei metro system and determine which stations are more central. This information is useful to analyze the vulnerability of the metro system that considers the recovering ability after system suffers a disaster. Finally, the comparisons among degree centrality, closeness centrality, betweenness centrality, and eigenvector centrality are made to show their differences and characteristics.**

Keywords—graph Fourier transform, network centrality, metro system, graph signal processing, vulnerability

I. INTRODUCTION

In recent years, graph signal processing (GSP) has received more and more attentions in many application areas including social network, sensor network, transportation network and biological network [1][2]. In the GSP, the irregular graph signals are represented by vertices, edges, weights and signal values over the vertices, so conventional regular time-domain signal processing methods can not be applied to process these irregular graph structured signals. To address this problem, the graph Fourier transform (GFT), graph filter, graph convolution and graph down-sampling operation have been developed to process graph signals. In [3], graph signal de-noising problem is solved by low-pass graph filtering method. In [4], missing temperature data recovery algorithm is presented by using the band-limited property of the signal in GFT domain. In [5], the brain signal is analyzed by using spectral GFT method. In [6], the graph signal processing method is developed by using orthogonal graph filter bank. In [7], the graph convolutional neural networks are used to solve the classification problem of graph-structured data.

On the other hand, measure of the centrality of complex network is important for us to understand the structures and dynamics of network [8]. The well-known conventional centrality measures include betweenness centrality (BC), closeness centrality (CC), degree centrality (DC), eigenvector centrality (EC) and PageRank centrality (PRC). These measures are useful to identify the central nodes or recognize key individual. The BC measures how a node behaves like a bridge in the network. A node with high BC often connects different parts in the network. The CC measures how a node in the "middle" of the network. The CC center provides the smallest sum of distances from CC node to all of other nodes. The DC measures the influence of node importance. Moreover, the evaluation of EC is based on the idea that the most central node is connected to more powerful nodes. The PRC is a popular centrality metric to measure relative importance of a node within a node-set.

In [9], the graph Fourier transform in GSP has been used to define a new centrality metric that quantify how important a particular node to other nodes in a network. This metric is called as graph Fourier transform centrality (GFTC) which not only utilizes the local properties, but also considers the global properties of a network topology. The comparisons of GFTC with BC, CC, DC, EC and PRC have been made in the context of various real-world networks [9]. Due to the success of GFTC, it is interesting for us to use GFTC to analyze the centrality of transportation networks. In this paper, the GFTC analysis of the Taipei metro system will be investigated. This centrality is useful for us to assess the reliability and robustness of metro system. This paper is organized as follows. In section II, the analysis problem of Taipei metro system is described and four conventional centrality measures are briefly reviewed. In section III, the computation algorithm of the GFTC is presented and uses it to analyze Taipei metro system in order to determine which stations are more central. Finally, experimental results are shown and conclusions are made.

II. PROBLEM STATEMENT AND CENTRALITY MEASURE

The route map of Taipei metro system is shown in Fig.1 which can be obtained from the website in [10]. Based on this route map, an un-weighted undirected graph $G=\{V,E\}$ can be constructed and is depicted in Fig.2. The V is the node set $\{v_1,v_2,\cdots,v_N\}$ and E is the edge set $\{e_1,e_2,\cdots,e_M\}$. Each node corresponds to a metro station, and each edge corresponds to a link between two stations. For Taipei metro system, number of stations is $N=120$ and number of edges is $M=136$. To study the structures and dynamics of the metro system, the centrality measure of complex network is used to study the node importance and identify central stations of the Taipei metro system. Before the graph Fourier transform centrality is used as node importance measure, four conventional centrality measures of BC, CC, DC and EC are first reviewed below.

The betweenness centrality measure $BC(v_n)$ at station v_n of the metro system is defined as

$$BC(v_n)=\sum_{n\neq j\neq k}\frac{g_{jk}(v_n)}{g_{jk}} \tag{1}$$

where $g_{jk}(v_n)$ is the number of shortest paths between v_j and v_k that pass through station v_n, and g_{jk} is the total number of the shortest paths from v_j to v_k. The station with high betweenness centrality measure means that it connects different parts of the metro system and behaves like a "bridge" in the system. Fig.3 shows the color representation of the resultant betweenness centrality measure of the Taipei

978-1-7281-9362-5/20/$31.00 ©2020 IEEE
DOI 10.1109/IS3C50286.2020.00059

metro system. It can be seen that the Minquan W. Rd. station with label 17 has the highest the BC measure. This is because station v_{17} connects many parts in the system.

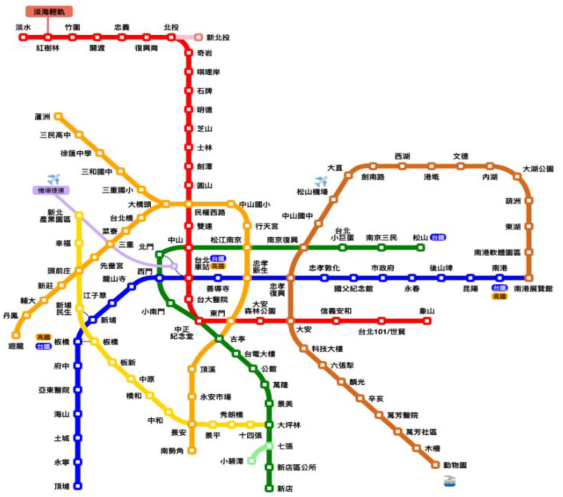


Fig. 1. The route map of Taipei metro system. There are 120 stations in this system.

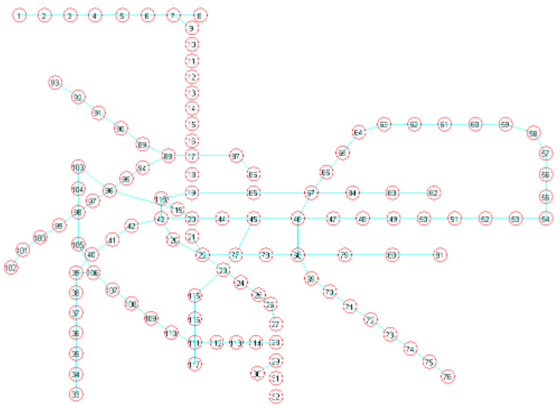


Fig. 2. The graphical representation of Taipei metro system. The number in each node is vertex label.

The closeness centrality (CC) measures how close one station is to other stations in the metro system. The closeness centrality measure $CC(v_n)$ at station v_n is defined as the inverse of the sum of length of the shortest path between v_n and all other stations in the metro system, that is,

$$CC(v_n) = \frac{1}{\sum_{i=1}^{N} d(v_n, v_i)} \tag{2}$$

where $d(v_n, v_i)$ is length of the shortest path between station v_n and v_i. The station with high closeness centrality measure means that it can quickly arrive at all other stations and behaves like a "middle" of metro system. Fig.4 shows the color representation of the resultant closeness centrality measure of the Taipei metro system. It can be observed that the Zhongxiao Xinsheng station with label 45 has the highest the CC measure. This is because station v_{45} locates near the middle part of the system.

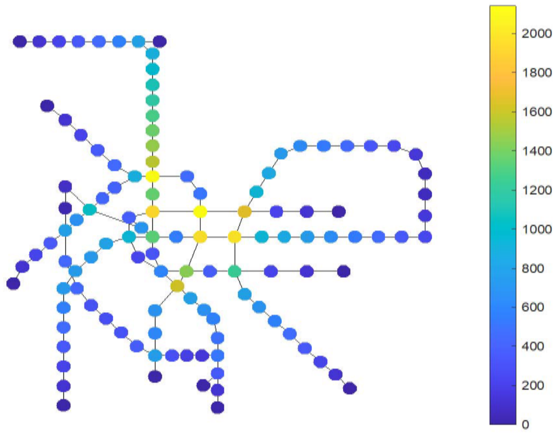


Fig. 3. The color representation of the betweenness centrality measure of Taipei metro system.

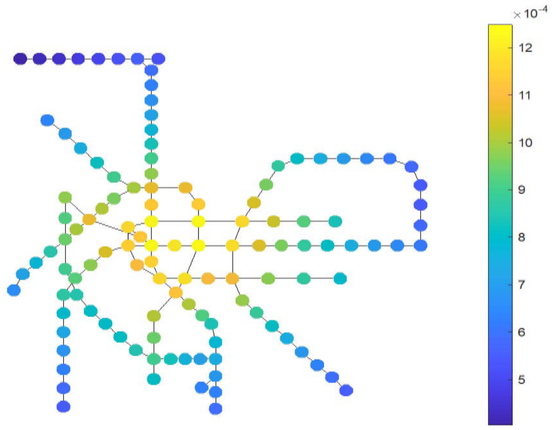


Fig. 4. The color representation of the closeness centrality measure of Taipei metro system.

The degree centrality (DC) is the number of the nearest neighbors that simply corresponds to degree, and clearly measures the ability of a station to connect directly with other stations. The degree centrality $DC(v_n)$ at station v_n is defined as

$$DC(v_n) = \sum_{m=1}^{N} a_{n,m} \tag{3}$$

where $a_{n,m} = 1$ if station v_n connects to station v_m, otherwise $a_{n,m} = 0$. The station v_i is more central than v_j if $DC(v_i)$ is greater than $DC(v_j)$. From the graph representation in Fig.2, it can be observed that the station v_{20} has the highest degree 5, so degree central station is Taipei main station with label 20.

The eigenvector centrality (EC) assigns importance of station v_n proportional to the importance of its neighbor stations. We can view EC measure as the steady state of each station passing its importance to all of its neighbors. Thus, the eigenvector centrality $EC(v_n)$ at station v_n is the n-th component of the eigenvector $\boldsymbol{q}$ associated with the largest eigenvalue λ of the adjacency matrix $\boldsymbol{A}$, that is,

$$EC(v_n) = \boldsymbol{q}(n) \quad (4)$$

Typically, we normalize $\boldsymbol{q}$ so that its Euclidean norm is one. For eigenvector centrality measure, the central station is also Taipei main station with label 20 if the graph representation in Fig.2 is used to compute the EC.

So far, conventional centrality measures have been briefly reviewed. In next section, the graph Fourier transform centrality (GFTC) will be studied.

III. GRAPH FOURIER TRANSFORM CENTRALITY

In this section, the definition of graph Fourier transform (GFT) is first reviewed briefly. Then, the importance signal of a reference station in metro system is defined. Finally, the computation method of GFTC is described.

A. Graph Fourier Transform

Let the $N \times N$ symmetric adjacency matrix of the un-weighted undirected graph $G = \{V, E\}$ be denoted by $\boldsymbol{A}$. And, the diagonal degree matrix of graph is defined as $\boldsymbol{D} = diag[d_1, d_2, \cdots, d_N]$, where d_n is the sum of all edge weights connected to vertex v_n. The graph Laplacian matrix is given by $\boldsymbol{L} = \boldsymbol{D} - \boldsymbol{A}$, where $\boldsymbol{I}$ is an identity matrix. The eigen-decomposition of the Laplacain matrix $\boldsymbol{L}$ is

$$\boldsymbol{L} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T \quad (5)$$

where matrix $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_N]$ is composed of the eigenvectors and diagonal matrix $\boldsymbol{\Lambda} = diag[\lambda_1, \lambda_2, \cdots, \lambda_N]$ is composed of the eigenvalues. The properties $\boldsymbol{U}\boldsymbol{U}^T = \boldsymbol{I}$ and $\lambda_1 \le \lambda_2 \le \cdots \le \lambda_N$ are satisfied in the above decomposition. A graph signal on G can be represented by the vector $\boldsymbol{f} = [f(1), f(2), \cdots, f(N)]^T$ whose the n-th component $f(n)$ is the value of the signal on vertex v_n. The signal $\boldsymbol{f}$ can be transformed into the spectral domain by the graph Fourier transform (GFT) $\hat{\boldsymbol{f}} = \boldsymbol{U}^T\boldsymbol{f}$. The vector form of GFT is denoted by $\hat{\boldsymbol{f}} = [\hat{f}(\lambda_1)\hat{f}(\lambda_2), \cdots, \hat{f}(\lambda_N)]^T$. The inverse graph Fourier transform (IGFT) is then is given by $\boldsymbol{f} = \boldsymbol{U}\hat{\boldsymbol{f}}$.

B. The Importance Signal

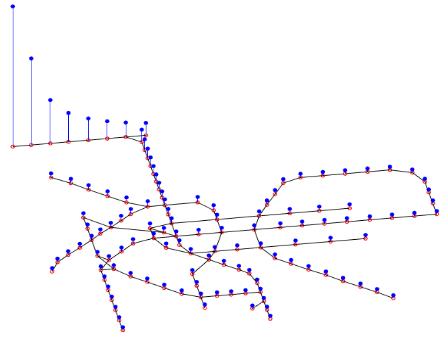
For a given reference station v_n in the metro system, its importance signal is a graph signal that is characterized by the inverse of the length of the shortest path from other station v_i to station v_n. The larger length to reach station v_n, the lower importance to the reference station v_n. Thus, the importance signal $\boldsymbol{f}_n$ corresponding to station v_n is given by the graph signal $\boldsymbol{f}_n = [f_n(1), f_n(2), \cdots, f_n(N)]^T$ where $f_n(i)$ is

$$f_n(i) = \begin{cases} \beta & if \quad i = n \\ \dfrac{\frac{1}{d(v_n, v_i)}}{Q_n} & if \quad i \ne n \end{cases} \quad (6)$$

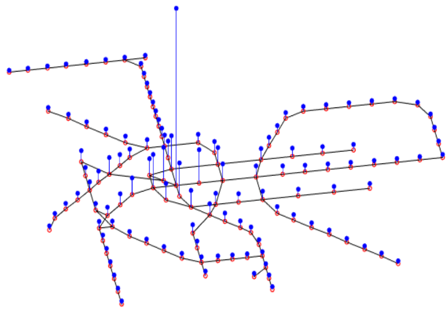
with

$$Q_n = \sum_{i=1, i\ne n}^{N} \frac{1}{d(v_n, v_i)} \quad (7)$$

In the above, $d(v_n, v_i)$ is the length of the shortest path from station v_n to v_i, and Q_n is used to normalize the signal such that the sum of signal values, except at station v_n, is equal to unity, that is, $\sum_{i=1, i\ne n}^{N} f_n(i) = 1$. The signal value $f_n(n)$ at station v_n is chosen as β, that is, $f_n(n) = \beta$. The variation in the importance signal $\boldsymbol{f}_n$ can be employed to measure the importance of the station v_n. The higher variation of the graph signal $\boldsymbol{f}_n$, the more importance of the station v_n. Fig.5(a) and (b) show the importance signals $\boldsymbol{f}_1$ and $\boldsymbol{f}_{20}$ with $\beta = 0.2$. It can be seen that the variation of signal $\boldsymbol{f}_{20}$ is higher than that of the signal $\boldsymbol{f}_1$. Thus, the Taipei main station v_{20} at the middle part of system is more important than Tamsui station v_1 at the periphery or border of system.



(a)



(b)

Fig. 5. The importance signals $\boldsymbol{f}_1$ at station v_1 and $\boldsymbol{f}_{20}$ at station v_{20} in the Taipei metro system. (a) Signal $\boldsymbol{f}_1$. (b) Signal $\boldsymbol{f}_{20}$.

C. Computation Method of GFTC

Before the GFTC of the metro system is defined, let us study the spectrum of the importance signal $\boldsymbol{f}_n$ at station v_n. The GFT spectrum of graph signal $\boldsymbol{f}_n$ is given by $\hat{\boldsymbol{f}}_n = \boldsymbol{U}^T\boldsymbol{f}_n$, where matrix $\boldsymbol{U}$ is eigenvector matrix of the Laplacian matrix $\boldsymbol{L}$ in (5). The transformed spectrum $\hat{\boldsymbol{f}}_n$ can be written as the vector form $\hat{\boldsymbol{f}}_n = [\hat{f}_n(\lambda_1)\,\hat{f}_n(\lambda_2),\cdots,\hat{f}_n(\lambda_N)]^T$. Fig.6(a) and (b) show the spectra $\hat{\boldsymbol{f}}_1$ and $\hat{\boldsymbol{f}}_{20}$ of the importance signals $\boldsymbol{f}_1$ and $\boldsymbol{f}_{20}$ depicted in Fig.5. It can be seen that the low-variation signal $\boldsymbol{f}_1$ has small high-frequency component, and the high-variation signal $\boldsymbol{f}_{20}$ has large high-frequency component. Thus, the importance information is encoded in the high-frequency component of spectrum. So, the importance measure $GFTC\ (v_n)$ at station v_n is defined as

$$GFTC\ (v_n) = I_n = \sum_{k=1}^{N} w(\lambda_k)\,|\,\hat{f}_n(\lambda_k)\,| \qquad (8)$$

where weights $w(\lambda_k)$ is chosen as a exponentially increasing function with frequency λ_k , that is,

$$w(\lambda_k) = e^{\alpha\lambda_k} - 1 \qquad (9)$$

In [9], the choice of parameter $\alpha = 0.1$ is suggested because it provides good experimental results. Because $\lambda_1 = 0$, we have $w(\lambda_1) = 0$, that is, zero-weight is assigned to the zero-frequency component. Moreover, from (9), it is clear that the large weight $w(\lambda_k)$ is assigned to high-frequency component for computing importance measure I_n because large-variation corresponds to the large importance. For the GFT spectra $\hat{\boldsymbol{f}}_1$ and $\hat{\boldsymbol{f}}_{20}$ in Fig.6, their importance measures are given by $I_1 = 0.05$ and $I_{20} = 0.43$. Clearly, I_{20} is greater than I_1, so station v_{20} is more important than station v_1.

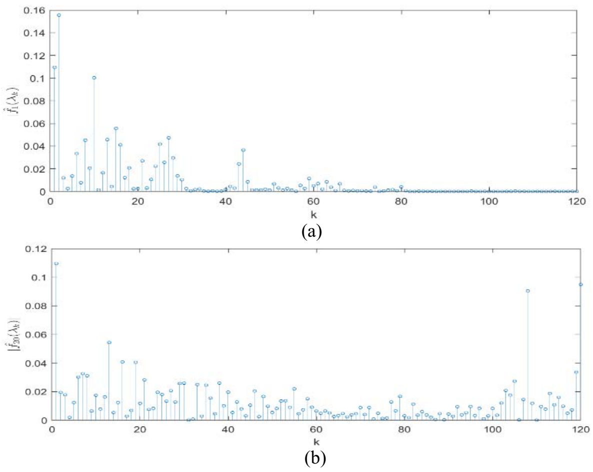


Fig. 6. The spectra $\hat{\boldsymbol{f}}_1$ and $\hat{\boldsymbol{f}}_{20}$ of the importance signals $\boldsymbol{f}_1$ and $\boldsymbol{f}_{20}$. (a) Spectrum $\hat{\boldsymbol{f}}_1$. (b) Spectrum $\hat{\boldsymbol{f}}_{20}$.

Fig.7 shows the color representation of the resultant graph Fourier transform centrality (GFTC) measure of the Taipei metro system with $\beta = 0.2$. It can be observed that the Songjiang Nanjing station with label 85 has the highest GFTC measure. Clearly, the center of GFTC is different from the centers of BC, CC, DC, and EC in section II. Because each centrality measure has its unique feature, so it is not easy to say which centrality measure is a better choice. The feature of GFT centrality is that it considers local and global characteristics of a node simultaneously, as described in [9]. Thus, GFT center locates near betweenness center, closeness center and degree center.

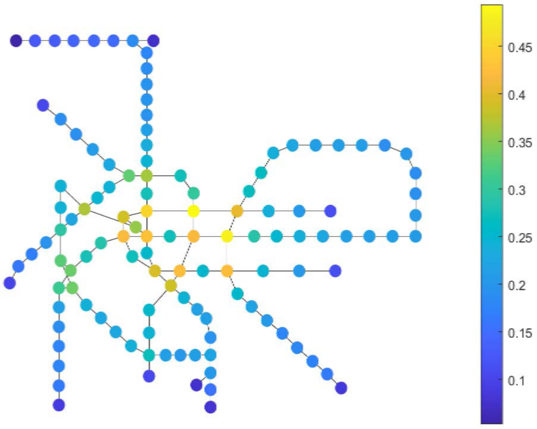


Fig. 7. The color representation of the graph Fourier transform centrality (GFTC) measure of Taipei metro system.

IV. CONCLUSIONS

In this paper, the GFTC for Taipei metro system has been studied. Three steps to compute GFTC is summarized below: First, the importance signal is calculated from the shortest paths among all source and destination station pairs in the metro system. Second, the values of GFT of importance signal corresponding to each reference station is computed. Third, the GFTC is obtained by weighted average the absolute values of GFT spectrum coefficients. However, only GFTC of Taipei metro system is studied here. Thus, it is interesting to study the GFTC of other metro systems of big cities in the world.

REFERENCES

[1] P.M. Djuric and C. Richard, Coorperative and Graph Signal Processing: Principle and Applications, Academic Press, 2018.

[2] A. Ortega, P. Frossard, J. Kovacevic, M.F. Moura and P. Vandergheynst, "Graph signal processing: overview, challenges, and applications" *Proceedings of the IEEE*, vol.106, no.5, pp.808-826, May 2018.

[3] W. Waheed and D.B.H. Tay, "Graph polynomial filter for signal denoising," *IET Signal Process.*, vol.12, Iss.3, pp.301-309, 2018.

[4] C.C. Tseng, S.L. Lee and R.H Su, "A missing temperature data estimation method using graph Fourier transform," *IEEE ICCE-TW*, pp.87-88, June 2017.

[5] W. Huang, L. Goldsberry, N.F. Wymbs, S.T. Grafton, D.S. Bassett and A. Ribeiro, "Graph frequency analysis of brain signals", *IEEE Journal of Selected Topics in Signal Process.*, vol.10, no.7, Oct. 2016.

[6] C.C. Tseng and S.L. Lee, "Design of orthogonal graph filter bank with known eigenvalues of Laplacian matrix", *IET Signal Process.*, vol.13, iss. 5, pp.551-561, 2019.

[7] R. Levie, F. Monti, X. Bresson and M.M. Bronstein, "CayleyNets: Graph convolutional neural networks with complex rational spectral filters," *IEEE Trans. Signal Process.*, vol.67, no.1, pp.97-109, Jan. 2019.

[8] B.S. Manoj, A. Chakraborty and R. Singh, Complex Networks: A Networking and Signal Processing Perspective, Prentic-Hall, 2018.

[9] R. Singh, A. Chakraborty and B.S. Manoj, "GFT centrality: A new node importance measure for complex networks," Physica A, vol.487, pp.185-195, 2017.

[10] https://piliapp.com/mrt-taiwan/taipei/.