PHYSA: 16793 Model 3Gsc pp. 1–8 (col. fig: NIL)

ARTICLE IN PRESS

Physica A xx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Analysis of community properties and node properties to understand the structure of the bus transport network

Yeran Sun [a,*], Lucy Mburu [a], Shaohua Wang [b]

[a] GIScience Research Group, Institute of Geography, Heidelberg University, Berliner Strasse 48, Heidelberg 69120, Germany

[b] Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Datun Road 11A, Beijing 100101, China

HIGHLIGHTS

- We analyze the community structure of the bus network and identify the important nodes in the network.
- The intra-urban bus network has a multi-community structure.
- The geographic characteristics of communities somewhat reflect the socio-economic division in the city.
- The majority of the important nodes (hubs) are clustered in the city center, implying the majority of bus lines are likely to go through the city center.

ARTICLE INFO

Article history:
Received 13 May 2015
Received in revised form 22 November 2015
Available online xxxx

Keywords:
Bus networks
Community structure
Node prominence
Geographic characteristics
Network topology

ABSTRACT

Akin to most infrastructures, intraurban bus networks are large and highly complex. Understanding the composition of such networks requires an intricate decomposition of the network into modules, taking into account the manner in which network links are distributed among the nodes. There exists for each set of highly interlinked nodes little connectivity with the next set of highly interlinked nodes. This inherent property of nodes makes community detection a popular approach for analyzing the structure of complex networks. In this study, we attempt to understand the structure of the intraurban bus network of Ireland's capital city, Dublin in a two-step approach. We first analyze the modular structure of the network by identifying potential communities. Secondly, we assess the prominence of each network node by examining the module-based topological properties of the nodes. Results of this empirical study reveal a clear pattern of independent communities, indicating thus, an implicit multi-community structure of the intraurban bus network. Examination of the geographic characteristics of the identified communities shows a degree of socio-economic divisions of the Dublin city. Furthermore, a large majority of the important nodes (vital transportation hubs) are located at the city center, implying that most of the bus lines in Dublin city tend to intersect the city's core.

1. Introduction

A clear knowledge of the structure and properties of public transportation networks is crucial for the urban planning and administration, policy enforcement, and disaster management [1]. The literature suggests that urban infrastructures, such as the bus, subway, railway, and aircraft networks have inherent small-world properties [2], and a complex topological

* Correspondence to: Room 101, Berliner Strasse 48, Heidelberg 69120, Germany. Tel.: +49 6221 54 5528; fax: +49 6221 54 4529.
E-mail address: yeran.sun@geog.uni-heidelberg.de (Y. Sun).

http://dx.doi.org/10.1016/j.physa.2015.12.150
0378-4371/

Please cite this article in press as: Y. Sun, et al., Analysis of community properties and node properties to understand the structure of the bus transport network, Physica A (2016), http://dx.doi.org/10.1016/j.physa.2015.12.150

structure [3–11]. Bus networks are the most popular kind of intraurban public transportation networks in many cities, Q3 and previous research has strived to understand the statistical mechanics of these networks [1,8]. Exploratory research has led to the discovery of three coordinated properties that characterize the bus networks as complex: (1) the degree distribution of nodes follows a power law or an exponential law, (2) the relationship between the degree of nodes and their relative influence is positive and linear, and (3) bus networks exhibit a high clustering coefficient, suggesting that such networks are 'small-worlds' [1,8,12–14]. Complex networks exhibit an unequal degree distribution, and this distribution forms the basis for dividing the network into smaller clusters. A "*community*" is, thus, detected from nodes with strong internal connections and weak connections between different clusters, and each community corresponds to a network component or sub-network [15]. This study, therefore, applies a community-based approach to analyze different properties of the network structure.

The purpose of this study is to investigate the structure of an intraurban bus network. Specifically, we assess the network community structure and subsequently analyze the roles of nodes in the bus network. The remainder of this paper is organized as follows. Section 2 introduces the approach to analyzing the network structure, while Section 3 presents the empirical study and the relevant results, and a discussion of the results. We conclude with recommendations for future research.

2. Methodology

This section advances the approach adopted for analyzing the bus network structure. First we introduce the strategies for partitioning the network into communities and assessing the geographic characteristics and topological structure of the primary communities. Afterwards, the section presents the application of module-based measures to describe the role of nodes in the network. The advanced approach will be applied in the subsequent sections to identify and analyze important bus network nodes.

2.1. Community detection

Modularity-based measures have been proposed for community detection [16–18]. Given a partition of a complex network into modules (sub-networks), the network modularity, M quantifies the strength of the division as:

$$M = \sum_{j \in J} \left[\frac{l_j}{L} - \left(\frac{d_j}{2L} \right)^2 \right] \tag{1}$$

where L is the number of network links, l_j is the number of links between nodes of module j (i.e., the number of intra-module links within module j), d_j is the sum of degrees of all nodes *within* module j, and $2L$ is the sum of the degrees of all nodes in the network. J is the set of modules divided. This definition of modularity can be justified. An accurate partition of the network into modules is designed to maximize the links *within* modules and minimize the links *between* modules. Accordingly, a high value of M corresponds to a solid network partition. However, if the sole intention is to minimize the number of between-module links (or, equivalently, to maximize the number of within-module links) the optimal partition will consist of a single module and no between-module links. Eq. (1) alleviates this problem by imposing the constraint that M is zero if the nodes are randomly located across the modules, or if all nodes belong to the same cluster [16,17].

Several modularity-based community detection algorithms have been proposed for partitioning complex networks, including FastGreedy [19], Spinglass [20], Walktrap [21], and Infomap [18]. A recent study by Ref. [15] comparing the community detection algorithms illustrated that Infomap outperforms the other widely used algorithms in its efficiency. This discovery guides the choice of the Infomap algorithm for community detection in the current study. Infomap detects the inherent community structure by minimizing the expected description length of a random walker's trajectory [18]. The algorithm employs the probability flow of random walks along a network as a proxy for information flow in the real system, and thus, it decomposes the network into modules by compressing a description of the probability flow [18].

2.2. Assessing the influence of network nodes

Two indicators are popular for measuring a node's influence within the network. The *within-module degree z*-score measures the degree centrality within a module, and the *participation coefficient* (p-value) measures the amount of variation in intermodular node connections [7]. The rationale for these indicators is that nodes which possess similar topological properties have high probability for playing similar roles. Within the framework of a complex network, the role of a certain node, v_i is defined by its location with reference to other nodes belonging to its module, and also by how well it connects to the nodes of other modules [7].

The modules of a complex network are structurally distinct in their organization, and they range from completely centralized modules (in which a few central nodes are linked to all other nodes) to completely decentralized (all nodes are similarly connected). It follows naturally, therefore, that nodes with similar roles will possess a relatively similar within-module connectivity. On the one hand, the z-score quantifies the degree of connectivity of a certain node, v_i to the other

Please cite this article in press as: Y. Sun, et al., Analysis of community properties and node properties to understand the structure of the bus transport network, Physica A (2016), http://dx.doi.org/10.1016/j.physa.2015.12.150