Annals of the American Association of Geographers
Routledge
Taylor & Francis Group

Mesoscale Structures in World City Networks

Journal:	*Annals of the American Association of Geographers*
Manuscript ID	AN-2017-0634.R3
Manuscript Type:	Regular Manuscript
Manuscript Category:	Geographic Methods
Key Words:	World City Network, Weighted Stochastic Block Model (WSBM), Imaginaries of globalization, Mesoscale Analysis, Webpage Big Data

Mesoscale Structures in World City Networks

Wenjia Zhang

(Corresponding Author)

School of Urban Planning & Design

Peking University Shenzhen Graduate School

Shenzhen, Guangdong 518055

China

Email: zhangwj@pkusz.edu.cn

Jean-Claude Thill

Department of Geography & Earth Sciences

University of North Carolina at Charlotte

Charlotte, NC 28221

USA

Email: jean-claude.thill@uncc.edu

Abstract: This study conceptualizes variant mesoscale structures (e.g., core-periphery, community, flat-world, or hybrid) in world city networks (WCNs) that enable us to understand the grouping features of cities with similar roles and positions, which are defined by distinctive relational patterns between cities as well as groups. A new analytical framework relying on a Bayesian-inference weighted stochastic block model (WSBM) is proposed to infer and compare latent mesoscale structures. This framework is superior to existing city-clustering approaches – which either *ex ante* postulates a mesoscale structure before clustering (e.g., the community detection method) or *ex post* explains the roles and positions of cities after clustering based on their similarities in attributes – and thus avoids the traps of *methodological determinism and territorialism*. For substantiating mesoscale structures, we study a WCN of 126 cities, between which the relational strength is measured by the pairwise co-reference frequency of cities on massive Internet webpages collected by a webometrics approach. Modeling results show that the WCN reflected on webpages has a distinctive multi-cores-peripheries structure mixed with communities. We also develop a likelihood-based estimation to compare alternative mesoscale structures and find that the WCN configuration differs from many geographical imaginaries of globalization, such as a homogenous world, a ranking hierarchy, or a territorialist regional geography. This study asks for more comparative analyses of mesoscale structures in the future research agenda of macroregional or world city network studies.

Key Words: *Imaginaries of globalization, Mesoscale Analysis, Webpage Big Data, Weighted Stochastic Block Model (WSBM), World City Networks*

Partaking in globalization, many cities are interconnected economically, politically, environmentally, and culturally and have become critical nodes embedded in world city networks (WCNs). A WCN generally consists of a collection of city nodes and a set of links between them depicting the existence or strength of their functional connections. Decades of research have documented such intercity relations in the form of co-reference of cities in business news (Pred 1980; Taylor 1997), air flight connection (Smith and Timberlake 1995; Derudder and Witlox 2005), co-location of advanced producer services (APS) firms (Beaverstock, Smith and Taylor 2000; Beaverstock et al. 2000; Taylor 2004), co-location of multinational corporate offices (Godfrey and Zhou 1999; Alderson and Beckfield 2004), and multiple relationalities linked by firm and non-firm actors (Coe et al. 2010; Mans 2014). While looking at specific intercity relations can reveal how distinctively cities are plugged into the global economy (Beaverstock, Smith and Taylor 1999; Taylor 2005; Derudder et al. 2008), other research has started to investigate the arrangements and distribution of power and control of world cities as a whole, that is, the structures embedded in WCNs (Wall and Van der Knaap 2011; Taylor et al. 2013). Yet, the identification and measurement of these structures remain a challenge to unravel, even though there are contrasted geographical imaginaries on what our living world looks like, such as a hierarchical world (Wallerstein, 1974; Friedmann 1986), a conflicted or territorialist civilization (Huntington 1997), a flat world (Friedman 2005), a spiky world (Florida 2008), or an overlapping world regionalization (Taylor et al. 2013).

This study categorizes WCN structures into three scales: microscale (or local), macroscale (or global), and mesoscale (or intermediate) in between, by reference to the body of literature in social network analysis (SNA) and network science (Rombach et al. 2014; Tunç and Verma 2015;

Jeub et al. 2015). Following this taxonomy, most WCN studies focus on microscale structures – local properties (e.g., attributes, interlocked connectivity, or centralities) of city nodes and dyads – or macroscale structures, such as city rankings and distributional 'laws' of local properties. Conversely, limited effort has been placed so far on mesoscale structures. These structures encompass the grouping of city nodes similarly embedded in a WCN based on distinctive *intercity interaction patterns* that define their roles and positions.

The tracing of mesoscale structures involves approaches that either *ex ante* postulate a mesoscale structure before clustering or *ex post* explain the roles and positions of cities after clustering based on their attribute similarities. To this end, three main perspectives can be followed. First, the most common *ex post approach* entails statistical clustering based on local attributes of city nodes and dyads, such as principal components analysis and hierarchical clustering analysis (Taylor and Walker 2001; Derudder et al. 2003; Liu et al. 2014). This approach may effectively identify groups of similar cities but neglects to compare the relational patterns between clusters. As a second approach, blockmodel has been specifically used to detect core-periphery (CP) structures (e.g., Alderson and Beckfield 2004). A third and more recent advance is the application of community detection methods to uncover city clusters (i.e., communities) such that cities are densely connected internally and sparsely linked externally (Neal 2014; Martinus and Sigler 2017; Rozenblat et al. 2017). The latter two approaches explicitly consider intercity relations but are typically *ex ante*, since both pre-assume there exists either a CP or a community structure within a WCN and exclude other potential structures. These *ex ante* approaches have an inherent deficiency defined here as *methodological determinism*, in which the methods used for city clustering will pre-determine the mesoscale structure in WCNs, probably conceal the "true" and hidden ones, and lead to a biased understanding of the role and position of world cities.

To overcome these analytical limitations, we develop a novel conceptual framework for mesoscale analysis of WCNs that identifies and quantifies potential mesoscale structures in terms of relational patterns between cities as well as groups. This enables us to underscore whether the global urban system ascribes to one or another theoretical framework, such as a CP hierarchy, a community structure implying territorialism or world regionalization, a flat-world 'blanket', or even a hybrid structure. Also, looking at variant mesoscale structures in WCNs aims to provide a new perspective, or even an analytic paradigm shift, in searching for the presence and configuration of hierarchies, territorialism and hybrids *in* networks, as a point of departure from the dualism tradition of hierarchy *versus* network.

This paper further puts forth a statistical inference method developed in network science, that is, the weighted stochastic block model (WSBM, Aicher et al. 2014), to detect stochastic, hybrid mesoscale structures in the relational properties of WCNs. In our empirical analysis, the linkage strength between two cities is taken to be the number of webpages on the massive information space of the World Wide Web (WWW) where pairs of city names are co-referred. This is akin to the seminal work of Pred (1980) and Taylor (1997), where city links are based on business news postings. The underlying assumption is that if two cities have more connections, they would be more frequently referenced on the same webpages. We identified 126 world cities based on existing literature and collected 28 sets of city network data using Google-based web crawling for one month in 2016. This study is among the first to systematically conceptualize and quantify multiple mesoscale structures in WCNs and provide a likelihood-based approach to comparing variant globalization imaginaries.

This article is organized as follows. The following section provides a literature review on structures and measures in WCN studies with a focus on the deficiency of local and global analysis, the conceptualization of mesoscale structure, and the dualism of hierarchy versus network. Next, it reviews existing methods for mesoscale analysis and presents the principle and features of the WSBM, and then describes the empirical analysis of the corpus of massive webpage documents. The subsequent section reports the analytical results. The last section discusses our results and concludes the research.

Structures in WCNs

Local and Global Structures in WCNs: A Need beyond Rankings

WCN structures are mainly investigated at the micro- and macro-scales. At the former, early studies have measured the power and position of specific cities based on various city-level aggregated attributes, such as population, economic product, and income level (Taylor and Derudder 2015). The macroscale approach explores the ranking and distribution of these attributes among a set of cities in order to detect uneven development in the global urban system. For example, many studies in urban economics and geography have found that the size distribution of cities fits a power law, also denoted as Zipf's law (Berry 1961). Many national urban systems fit this rank-size rule, implying a hierarchical sorting of cities where a large share of population is concentrated in a very small number of cities. Urban studies also provide comprehensive surveys on the control functions of world cities in the global economy, heavily relying on rankings of city attributes (Hall 1966; Hymer 1972). As noted by several WCN studies (Beaverstock, Smith and

Taylor 2000; Alderson and Beckfield 2004; Taylor and Derudder 2004), the structures found by the attribute-based approach overlook the control and power roles of world cities embedded in the external relations to other cities in the network. Subsequent globalization research has thus turned to emphasize the inter-city relational and networked essence of world cities, following three seminal conceptual contributions from John Friedmann's (1986) world city hypothesis, Saskia Sassen's (1991) global cities, and Manuel Castells' (1996) network society.

With the increasing availability of data on intercity linkages, WCN studies have developed new measures that substitute relational data to attribute data (Beaverstock, Smith and Taylor 2000). Such efforts are spearheaded by the Globalization and World Cities (GaWC) research group, with a particular emphasis on the co-location relationships of advanced producer service (APS) firms (Beaverstock, Smith and Taylor 2000; Taylor 2004), while some research follows the tradition of world system theory and SNA (Smith and Timberlake 1995; Alderson and Beckfield 2004; Neal 2012a). These studies mainly introduce three types of local-level analysis. The first measures the importance and central roles of specific city nodes in the WCN. For example, the GaWC group developed a measure of 'global network connectivity (GNC)' of a city, which is the sum of the interlocked links connecting with this city (Taylor 2001) or the well-known interlocking network model (INM, Derudder and Taylor, 2017). This connectivity measure is similar to the (weighted) degree centrality in SNA terminology (Taylor and Derudder 2015). Other centrality measures are also introduced, such as closeness, betweenness, and eigenvector centralities (Alderson and Beckfield 2004; Mahutga et al. 2010; Neal 2015; Martinus et al. 2015). Second, some studies have extended the analysis from city nodes to dyads, with a focus on a city's relations to all other cities in the WCN, also defined as 'hinterworld' (Rossi et al. 2007).

The third strand of local analysis is relatively qualitative, investigating ego-centric networks of specific cities (Beaverstock, Smith and Taylor 2000; Mans 2014).

As for the global properties of WCNs, most studies continue to apply city rankings, as earlier attribute-based studies did, but now sorting cities by relation-based local measures (Martinus et al. 2015; Derudder and Taylor 2017). Rankings have become a key tool for urban geography comparisons (Taylor 1997) and possibly the most popular approach to measuring the relative standing of cities in WCNs. However, this approach has many inherent drawbacks. First, a ranking often returns a list of top cities, ignoring bottom-tier cities. Second, the application of city rankings presents a multidimensionality issue. While diverse urban rankings can be created on multiple local measures, there still is no consensus on criteria that produce more robust rankings. It is thus difficult to compare different sets of findings and collapse them for a systematic understanding. Third, rankings are often regarded as a compromise strategy for classifying cities into groups. As Friedmann (1995: 23) argues, since there are no clear standards for "assigning particular cities to a specific place in the global system", "a rough notion" of city ranking "without further specification is all that is needed". This statement is criticized by some WCN scholars (e.g., Taylor and Derudder 2015) but, to some extent, it may explain the rationality behind the application of city ranking in WCN studies.

Fourth, while city rankings are often regarded as a measurement tool implying hierarchies of the world city systems (e.g., Friedmann 1986; Taylor 2009), they do not reveal the actual structure of hierarchies. Hierarchies are not a requirement of rankings. "A hierarchy cannot be simply inferred from a ranking of cities; there have to be some form of power relations between the cities that sort them into a hierarchy" (Taylor and Derudder 2015: 17). A world city hierarchy relying on node-based or dyad-based rankings is thus a problematic construct. According to traditional world system studies (Lloyd et al. 2009), an explicit hierarchy should look like a CP structure, which can only be identified through the relational patterns between core and periphery groups at the mesoscale level. Simply relying on city rankings cannot convince us which cities belong to the core or the periphery. Consequently, to unravel structures (including hierarchies) in WCNs, we need to examine the mesoscale properties and structures in WCNs beyond simple local measures and rankings.

Conceptualizing *Variant* Mesoscale Structures in WCNs: Regionalization, Core-Periphery, Flat-World, or Hybrid?

Mesoscale structure is a terminology from the field of network science (Rombach 2013; Tunç and Verma 2015). It describes the grouping properties that may not be apparent at either local or global scale. A mesoscale structure in a WCN can be defined as a partition of city nodes into groups based on their distinctive intercity interaction patterns, that is, equivalent social positions (Wasserman and Faust 1994). In other words, what defines a mesoscale structure is the presence of strong relational ties and similar positions (compared to the rest of the world), instead of common individual features of the city nodes (e.g., economic size, population, level of development) or the strength of their links. However, there is no consensus on mesoscale structures in WCNs and they are largely understudied compared to local and global structures. Borrowing similar concepts from SNA and network science (e.g., Faust and Wasserman 1992; Rombach et al. 2013; Aicher, Jacobs and Clauset 2014; Jeub et al. 2015), this article advances this body of literature with a conceptualization of potential mesoscale structures in WCNs, as shown in Figure 1.

[Figure 1 Inserted Here]

Figures 1A-D enumerate four possible mesoscale structures in a WCN partitioned in two clusters. They show the city-by-city adjacency matrix by block with darker colors representing stronger connections on average. Suppose a four-block adjacency matrix is denoted by $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$, and A_{ij} $(i, j = 1 \; or \; 2)$ is the average linkage strength between any two cities in blocks *i* and *j*. We can list four types of mesoscale structures by comparing the intra- and inter-block relations A_{ij}.

The first mesoscale structure is an (assortative) community (Figure 1A), which constructs a set of city nodes with denser linkages internally than with the rest of cities in the WCN (i.e., $A_{11} \approx A_{22} \gg A_{12}, A_{21}$). This community structure may exist in a real WCN at a very limited level of globalization, implying territorialism (nationalization or regionalization), in which intercity connections are strong within countries/regions, while relatively weak between countries/regions. In fact, many early studies of national urban systems and studies of globalization postulated such a territorialist structure by overlooking the connections external to the nation in their analysis (Pred 1977). These studies are often criticized as "methodological nationalism" (Wimmer and Schiller 2002; Beck 2007) or "methodological territorialism" (Brenner 1999). Instead, such community structure may reflect a true division of regional geographies in the world. Many studies highlight strong regional patterns among world cities (Taylor and Walker 2001; Derudder et al. 2003), while a recent study emphasizes that the boundaries of regionalization of cities are fuzzy and overlapping, rather than crisp and exclusive (Taylor et al. 2013). They typically rely on statistical clustering methods, such as principal components analysis, fuzzy set analysis, and hierarchical clustering analysis. Whereas community features may be detected in city networks, associated territorial properties are questionable because they fail to represent relational patterns between communities.

Some recent studies have directly investigated community structures in WCNs relying on community detection algorithms (see detailed reviews of algorithms in Fortunato 2010; Fortunato and Hric 2016). For example, Neal (2014) applied the widely-used Louvain method (a greedy algorithm for modularity maximization) to detect communities in the US air traffic network and found three communities in the city networks driven by either economic or spatial-proximity functions. Rozenblat et al. (2017) compared several modularity-maximization algorithms and selected the spin glass clustering algorithm to find communities in a large WCN defined by the co-location of various types of multinational firms. Their findings show multipolar-regionalization structures in the corporate WCNs. Martinus and Sigler (2017) used another modularity-maximization algorithm proposed by Clauset et al. (2004) to detect community structures in the city networks connected by Australia-based firms in four types of industries. They found significant community structures in different WCNs and also that world cities in the same communities have strong ties due to both spatial (i.e., geographic) proximity and non-spatial proximities (e.g., similar organizational norms, social culture, institution, and cognitive knowledge base). However, most community-detection methods postulate the existence of communities in a network and exclude other possible mesoscale structures. These studies could be easily trapped into the "*methodological determinism*" as defined previously. For instance, a CP structure (or another) could never be found by most community detection methods because their purpose is recognizing communities.

The second structure is the CP hierarchy (Figure 1B), in which city nodes are strongly connected both within the core and to a set of peripheral nodes that interact very little among themselves (i.e., $A_{11} > A_{12}, A_{21} \gg A_{22}$). Early studies mainly used the ranking technique to define top city groups. Influential work by Beaverstock, Smith and Taylor (1999) ranked a list of world cities based on the presence of four types of APS headquarters and categorized them into three groups: alpha, beta, gamma cities. Some recent WCN studies espouse the SNA approach and categorize cities into groups by their rankings on a variety of city centrality measures (Martinus et al. 2015; Martinus and Tonts 2015; Derudder and Taylor, 2017) or by clique detection (Derudder and Taylor 2005). In fact, grouping from centrality ranking is equivalent to finding cohesive sub-networks of cities that are mutually better connected, closer to each other (via reachability), or more densely connected than those in the rest of the network (Wasserman and Faust 1994). Corresponding techniques, such as *k*-cores, *k*-cliques, *k*-components, *k*-clan, *k*-club, and *k*-plex (Csermely et al. 2013), search for optimal sub-networks with *k* as a threshold to screen nodes and/or edges based on some centrality or linkage properties. The setting of the grouping boundaries thus boils down to delineating a set of *k*'s. This genre of analysis can partially reveal the hierarchical structure since it identifies some locally dense cores but may neglect the structure of other nodes/edges that are not in the cohesive subgroups. Also, it is difficult to settle on consistent criteria to determine the *k*'s for grouping (Csermely et al. 2013), leading to relatively subjective results. For example, it is difficult to articulate why there are 6 core primary cities or 10 alpha world cities, rather than any other numbers of groups, in Beaverstock, Smith and Taylor's (1999) study. It lacks of quantitative basis to differentiate groups by their importance or which ones play the core, periphery, or other roles.

The empirical study of Alderson and Beckfield (2004) stands out in this body of literature, as it retains the methodological tradition of world system research by using blockmodeling analysis (Snyder and Kick 1979; Smith and Timberlake 1995; Lloyd et al. 2009; Kick et al. 2011). The authors presume that a CP structure exists in a WCN and detect such a structure based on a renumbering process that discriminates city nodes into core and periphery clusters and determines the relations between these clusters (Borgatti and Everett 1999). Although the advantage of blockmodels is recognized that "the configurations of blocks are characterized specifically in relations between and within blocks" (Taylor 2006, 887), few studies have applied such an analytical tool. Recently, it has been proposed to bring back the world-systems tradition to the study of WCN by integrating Global Commodity Chains (GCC) studies, with a focus on disclosing the CP structure in WCNs (Brown et al. 2010). But such discussion has remained conceptual, qualitative, and conjectural, and lacks appropriate methodologies for detecting CP structures. In contrast, some scholars doubt whether a CP structure still exists in current global systems. In response to Brown et al.'s (2010) claim, Coe et al. (2010, 140) argued that "[s]uch a (structuralist) reading of the geographies of power [i.e., the CP structure, added by the authors], however, implies a Braudelian world of nested scalar hierarchies, which, arguably, no longer exists". It is worth noting that Coe et al. mainly questioned the necessity of refocusing on the CP dichotomy, particularly when intercity relations become dynamic and pluralistic. Also, the CP-detection methods have similar limitations due to the "methodological determinism" as community detection algorithms since they rule out other potential mesoscale structures.

The third mesoscale structure is that of a flat-world with insignificant difference in intra- and inter- block connections, i.e., $A_{11} \approx A_{12} \approx A_{21} \approx A_{22}$ (Figure 1C). In such structure, the world is flattened out and there are neither hierarchies nor communities, such that no cities have comparative advantage or disadvantage in their positions. This might be a world configuration imagined in Thomas Friedman's (2005) book *The World Is Flat*, which argues that the world becomes flat due to better connection and more cooperation globally. Another example is Wall and Van der Knaap's (2011) heterarchical (a complete or "universal") corporate network structures.

The fourth one is disassortative community or bipartite (or two-mode) structure (Figure 1D). A disassortative community is the inverse of a community structure, representing that nodes between groups are densely linked while those within a group are sparsely connected (i.e., $A_{12} \approx A_{21} \gg A_{11}, A_{22}$). It appears difficult to find disassortative properties in actual WCNs, but this structure may fit into two-mode networks with two sets of actors – for example, firms and cities – such as the interlocked city network defined by the GaWC group. The adjacency matrix of such a two-mode network looks like Figure 1D, in which the first part represents firm nodes while the second part includes all city nodes. In such a firm-by-city two-mode network, cities (or firms) have no direct connection with other cities (or firms) (i.e., A_{11}=0 and A_{22}=0), while firms are connected to cities (i.e., $A_{12} = A_{21} \gg 0$).

Actual WCN structures can be much more complex than the four-block network model presented above, especially when the number of partitions is large. Figures 1E-D show two typical hybrid mesoscale structures (Rombach et al. 2013), including a global community structure with regional cores and peripheries and a global core-periphery structure with regional communities. Conceivably, an increase in the number of blocks will raise the complexity of mesoscale structures in WCNs. An influential example is Friedmann's (1986) hierarchical classification of world cities that constructs a two-level city hierarchy in the world economy (i.e., primary versus secondary cities), each level of which has a CP structure (i.e., core versus semi-periphery cities), although no clear definitions were provided. While this classification scheme has reached great popularity, it remains to be validated and just provides a "possible" typology fitting into Friedmann's geographical image of the world economy. In addition, few studies have attempted to discover hybrid structures in WCN.

From the Dualism of Hierarchy *versus* Network to a Search for Hierarchies *in* Networks

Early literature on world cities retained the hierarchy thinking inherited from world-system theory (Smith and Timberlake 1995; Friedmann 1995; Taylor 1997; Beaverstock, Smith and Taylor 1999). Intercity-relation analysis has turned to underscore the network nature of WCNs by seemingly differentiating it from the traditional hierarchy thinking. For example, as Taylor states (2009), hierarchy represents verticality, a ranking space of competitive and powerful places, while network suggests horizontality, a flattened space of reciprocity and cooperation. Thus city rankings serve as a proxy for hierarchy and intercity competition "against the spirit of the GaWC project" (Taylor 2004). As reemphasized by Taylor (2011, 16), "our starting point is the specification of a world city network to replace the hierarchical theory in its various forms".

Such dualism may have been helpful early on when relational and network thinking was introduced in world cities research but, as argued here, it may have hindered further development of WCN theory and analysis. Treating networks as a construct superior to hierarchies may give researchers an excuse to overlook potential hierarchical structures in WCNs. For example, the dualism may falsely lead to regard non-hierarchical structures as a composite type of network structure without further conceptualization, seemingly nurturing the illusion that describing and visualizing the network properties at local and global scales can bring out results that contradict hierarchical analysis, like rankings.

Some recent studies appear to blur or modify the dichotomy between hierarchy and network. For example, the GaWC group advances an agenda to return to the analysis of "network with hierarchical tendencies" by assigning cities into ordering categories (Taylor et al. 2009) and to refocus on the CP structure (Blown et al. 2010), although no methodological improvement has been clearly articulated. Recent studies also have paid attentions to questions such as "how else would cities relate to each other except through hierarchies"? (Taylor 2011, 16). Related to it, Wall and Van der Knaap (2011) provided some morphological evidence of the "coexistence" of hierarchical (a star-shape network) and heterarchical (a complete network) city network structures based on simple network visualization. While these studies seem to recognize that the hierarchy is just one type of network structures, they deliver less theoretical or methodological discussions on what other mesoscale structures may possibly exist in WCNs. In addition, as Neal (2012b) states, while the WCN is likely to have different structural forms (simply like a ring, star, or chain), the dataset chosen for analysis (e.g., the interlocking WCN data) may pre-determine the structure of the WCN, i.e., the problem of "structural determinism". This problem may also hinder further exploration on a variety of mesoscale structures in WCN studies.

Given the multiplicity and complexity of mesoscale structures in WCNs (Figure 1), this article argues for a shift in analytical framework, from a view rooted in the dualism of hierarchy *versus* network to a search for the presence and configuration of hierarchies and hybrids *in* networks. The new analytical approach allows for tracing stochastics, dynamics, multiplicity, and hybrids of network properties and structures, rather than treating a WCN as fixed, static, singular, and uniform as in much of the existing GaWC research (Coe et al. 2010; Mans 2014). This new framework needs to develop methods for mesoscale analysis, instead of only focusing on local and global measures. The following section will thus review the approaches to detecting mesoscale structures and discuss how to apply them to study WCNs.

Methodology: A Weighted Stochastic Block Model

Blockmodeling is a well-developed method of positional and role analysis in SNA to identify prevailing mesoscale structures or image matrices without assuming any specific structure a priori (White et al. 1976; Faust and Wasserman 1992). It classifies nodes into clusters and determines inter-cluster relational bundles such that nodes within the same cluster are deemed to have equivalent positions (i.e., equivalence). There are mainly two equivalences: social actors in a network are said to be structurally equivalent when they link with identical neighbors, or regularly equivalent when they link with equivalent others (Wasserman and Faust 1994). For searching structurally-equivalent positions, blockmodeling analysis aims to detect two types of *ideal blocks* (1-blocks and 0-blocks) in an adjacency matrix. A more relaxed blockmodel allows for regular blocks (with at least one “1” in each row and/or column) to detect regularly-equivalent positions. These methods search for a partition to minimize the error score.

An error is added if a cell is 0 (or 1) but should be 1 (or zero) under the definition of ideal or regular blocks. After that, researchers can shrink the original network to a reductive network with blocks as nodes and then match the shrunk network with one of the predefined mesoscale structures (e.g., Figure 1). A major limitation is that the definitions of ideal blocks or equivalence are too restrictive such that it may filter out many possible mesoscale structures. For instance, the CP blocks defined in Borgatti and Everett (1999) are not fully identical to the ideal or regular blocks. Also, without a statistical test, this approach cannot tell us how well the network data fit into specific mesoscale structures.

To overcome these limitations, stochastic block models (SBMs) have been introduced by Holland et al. (1983). They first create a relaxed definition of equivalence, dubbed stochastic equivalence. Two nodes are stochastically equivalent if they have the same probability to connect to nodes in other groups. No predefined block structures are needed in SBM analysis; the distribution of ties between nodes only depends on the blocks to which the nodes belong. SBM has become a popular probabilistic model for learning mesoscale structures in networks. Early SBMs were applied to binary networks, while recent developments have extended them to consider weights and degree distributions (Karrer and Newman 2011; Aicher, Jacobs and Clauset 2014). This article mainly incorporates the settings of Aicher, Jacobs and Clauset's (2014) weighted SBM (WSBM).

Intuitively, the WSBM used in this article assumes a generative process of a weighted WCN that is formulated through two steps. First, each city node $i = 1, 2, \dots, N$ is assigned to a latent group membership $z_i \in \{1, \dots, K\}$. The integers N and K denote a given number of cities and groups, respectively, and z_i is the group index of node *i*. Second, city *i* builds a connection of weight a_{ij} to city *j*. Here a_{ij} is not deterministic but stochastic, distributed according to the appropriate edge parameter $\theta_{z_i z_j}$ that depends only on the group memberships of cities *i* and *j*. This setting reflects the principle of stochastic equivalence that all cities in a group have the same probabilistic connectivity (parameterized by $\theta_{z_i z_j}$) to the rest of the network. For example, if a_{ij} fits into a normal distribution, each edge bundle $(z_i z_j)$ is parameterized by mean and variance $\theta_{z_i z_j} = (\mu_{z_i, z_j}, \sigma^2_{z_i, z_j})$. Given a block matrix $\boldsymbol{\theta} = [\theta_{kk'}]_{K \times K}$ and a known distribution of $\theta_{z_i z_j}$, one can thus estimate a WCN of adjacency matrix $\boldsymbol{A} = [a_{ij}]_{N \times N}$.

The WSBM has a reversed inference process: given *K* and an observed WCN *G*, we can use the WSBM to infer the latent grouping $\boldsymbol{z}$ and stochastic block matrix $\boldsymbol{\theta}$. It thus becomes an optimization problem of likelihood maximum by choosing $\boldsymbol{z}$ and $\boldsymbol{\theta}$. The likelihood for a binary network of adjacency matrix *A*, when the edge weights are Bernoulli random variables (1 or 0), is given by (Holland et al. 1983):

$$P(\boldsymbol{A}|\boldsymbol{z}, \boldsymbol{\theta}) = \prod_{i,j} \exp\left(a_{ij} \log\left(\frac{\theta_{z_i z_j}}{1-\theta_{z_i z_j}}\right) + \log\left(1 - \theta_{z_i z_j}\right)\right) \tag{1}$$

Aicher et al. (2013) extends the SBM to weighted networks by associating θ with an exponential family of parametric distributions, including normal, exponential, Pareto, and Poisson distributions. This study selects the normal distribution as default. Suppose the mean and variance matrices are $\boldsymbol{\mu} = [\mu_{kk'}]_{K \times K}$ and $\boldsymbol{\sigma^2} = [\sigma^2_{kk'}]_{K \times K}$, the likelihood function becomes:

$$P(\boldsymbol{A}|\boldsymbol{z}, \boldsymbol{\theta}) = P(\boldsymbol{A}|\boldsymbol{z}, \boldsymbol{\mu}, \boldsymbol{\sigma^2}) = \prod_{i,j} \exp\left(a_{ij} \frac{\mu_{z_i, z_j}}{\sigma^2_{z_i, z_j}} - a^2_{ij} \frac{1}{2\sigma^2_{z_i, z_j}} - \frac{\mu^2_{z_i, z_j}}{2\sigma^2_{z_i, z_j}} - \log \sigma_{z_i, z_j}\right). \tag{2}$$

Aicher et al. (2014) apply a Bayesian approach to optimize the likelihood function for the WSBM by treating parameters $\boldsymbol{z}$ and $\boldsymbol{\theta}$ as random variables with an appropriate prior distribution $P(\boldsymbol{z},\boldsymbol{\theta})$. Based on Bayes' theorem, the posterior distribution $P(\boldsymbol{z},\boldsymbol{\theta}|\boldsymbol{A})$ is estimated as:

$$P(\boldsymbol{z},\boldsymbol{\theta}|\boldsymbol{A}) \propto P(\boldsymbol{A}|\boldsymbol{z},\boldsymbol{\theta})\,P(\boldsymbol{z},\boldsymbol{\theta}). \tag{3}$$

Aicher et al. (2014) develop a variational Bayes expectation-maximization algorithm to estimate $P(\boldsymbol{z},\boldsymbol{\theta}|\boldsymbol{A})$ using several approximation techniques and thus optimize an approximation of the likelihood function $P(\boldsymbol{A}|\boldsymbol{z},\boldsymbol{\theta})\,P(\boldsymbol{z},\boldsymbol{\theta})$ computed in Eq. (3). Simply put, the WSBM can search over various choices of parameters $\boldsymbol{z}$ and $\boldsymbol{\theta}$, detect the space of all partitions to find out the parameters with approximately maximal likelihood scores, and finally return the best partitions for identifying statistically significant mesoscale structures.

Mesoscale analysis in WCN studies can benefit from the WSBM method. First, WSBM is based on statistics and probability theory, providing statistical inference of mesoscale structures. Using the likelihood scores, it can readily determine the optimal number of partitions without an arbitrary guess and compare the goodness of partitions and infer whether a WCN falls into a territorialist geography (community), a CP hierarchy, or a hybrid. Second, it does not pre-define positions (like cores and peripheries) or block structures at the very beginning, as conventional blockmodels do. This allows for learning variant new mesoscale structures that are not presumed but might truly exist in some WCNs. Third, the method is particularly advantageous for weighted networks in this study, since it can avoid a conversion from weighted networks to binary networks. The conversion process could lead to significant information loss and distortion (Aicher et al. 2014). Fourth, many observed intercity relations are probabilistic; the stochastic nature of WSBM better allows for the sampling errors in the data collection process. Our study uses Matlab 2016a to run the WSBM algorithm and creates network visualizations using the network analysis package in Matlab and Gephi.

Case Study: A WCN Reflected in Massive Webpage Content

Intercity Relations Based on Webpage Content

The diversity of intercity relations has dramatically increased since the 1990s when a network-based research agenda emerged (Smith and Timberlake 1995; Taylor 1997; Beaverstock, Smith and Taylor 2000). Research mainly focuses on two ways to apprehend intercity relations, namely the corporate organization approach and the infrastructure approach (see reviews of Derudder (2008) and Taylor and Derudder (2015)). While both approaches emphasize the economic relations among cities, some criticisms have been placed at the "economism" of city linkages (Therborn 2011) and called for a "move beyond economic relations among cities and address the social, political, and cultural dimensions of the world city system" (Alderson and Beckfield 2006).

Differing from these two approaches, this study apprehends an intercity linkage as the co-occurrence frequency of two cities on the Internet, and specifically defines it as the number of webpages with both cities referenced. When two cities are more frequently referenced on webpages, they are assumed to be more tightly connected, including not only economic linkages but also transportation, tourism, historical, social, cultural, and governmental relationships, via different document contents (Liu et al. 2014; Salvini and Fabrikant 2015; Hu et al. 2017). Thus, the co-referencing linkage represents the cumulation of various intercity connections and serves as a measure of a city's aggregated external relations to the other cities in the WCN.

The approach of document analysis can be traced back to Pred (1980) and Taylor (1997), although their data source was business news. Beaverstock at al. (2000, 49) suggested that content analysis is a "surrogate measure of a city's external relations", providing "a continuous source of information on what a given editor thinks are the salient news stories of the day for a given readership". This approach has its own advantage (e.g., Beaverstock at al. 2000; Hoyler et al. 2010; Devriendt et al. 2011) because it provides an easy access to data showing dynamic (real-time or short-term) and varied intercity relations, and it is relatively easy to collect massive document data that cover the production and consumption of global information relevant to intercity linkages.

As actual cities are connected in multiplex, either perceptible or imperceptible networks, there are substantial limitations to any one set of relational data, even those used in the corporate or infrastructure approach. In addition, the structure of WCNs may be highly determined by the selection of city network datasets, i.e., the structural determinism (Neal 2012b). The document approach used here also has some measurement errors. First, it only records the relative importance of cities or city-dyads instead of their actual intercity relations. Second, there may be some subjectivity of documental data sources: one from the subjective editorial choice of contents relevant to their readers (e.g., types of news events) and the other from researchers' sampling selection of documental sources (e.g., the choice of newspapers). It is expected that these errors could be moderated when the sample size of documents is massive. Accordingly, compared to a relatively small amount of newspapers, webpage documents should be a better source for analyzing WCNs. By the end of 2016, there were about 50 billion webpages[1] and 49.2 percent of world population are Internet users[2]. Intercity relations extracted from massive webpages have a great potential to show an aggregation of multiplex connections among cities. Although several empirical studies have attempted to collect intercity connection data from webpages via search engines (e.g., Devriendt et al. 2011), their analytical approaches are similar to the GaWC studies that rely on local measures and city rankings, for example, the ranking of search counts.

Data Collection through Web Crawling

This study applies a webometrics approach (Almind and Ingwersen 1997; Thelwall 2009) to collecting intercity relational data. Ideally, a webometrics approach to WCN studies need to collect all public-access webpages from the WWW, identifying whether each webpage mentions city pairs, and compute the co-occurrence frequency of the city pairs. Given the extremely large number of webpages, this task is practically infeasible for individual researchers (cf. Hu et al. 2017). Thus, many webometrics studies have relied on search engines to obtain the numbers and trends of a certain keyword as a proxy for quantified indexes of influence and connection strength (Thelwall 2009). Many search engine companies, such as Google, Bing, and Yahoo Search, have served as a platform to crawl webpages periodically, store webpages with indexing, and retrieve the information from the webpages according to users's requests. Although the webpages indexed by search engines does not cover all webpages in the world, they represent the majority of webpages accessible to the public. By entering a keyword (e.g., two city names), we can obtain a list of best matching webpages, images, documents, and other electronic materials

[1] http://www.worldwidewebsize.com/
[2] http://www.internetworldstats.com/stats.htm

and the estimated total number of results found by the search engine.

This study uses the Advanced Google Search Engine as the platform to investigate the co-occurrence frequency of city pairs. When a pair of city keywords is entered into the search box, Google will return an estimated number (rounded to 1,000) of indexed webpages containing both city names. Compared to the idealistic approach, the Google search approach has several limitations.

First, most search engines are keyword-oriented such that they can return hit counts only when some keywords are entered. We thus need to select a list of cities as keywords in advance. In our study, we first create a pooled list of 327 cities that have been identified as world cities in at least one of 22 influential WCN studies in last 30 years. To further narrow down, only cities shown in at least two of these studies make it to our final list[3]. We finally selected 126 world cities, including 42 cities in Asia (33.3 percent), 41 in Europe (32.5 percent), 22 in North American (17.5 percent), 11 in South America (8.7 percent), 7 in Africa (5.6 percent), and 3 in Oceania (2.4 percent). It is worth noting that the selection of cities in many WCN studies is a major source of bias in analysis (Robinson 2002). For example, many WCN studies with a focus on APS and multinational companies often overemphasize the cities from the core economies of the United States and Europe but underestimate the less-developed countries. Also, the selection of cities and network data could pre-determine the structure observed in the WCNs (Neal 2012b). While the selection approach here may still replicate the selection biases of the 22 WCN studies, tapping into a literature contributed over a longer period can alleviate such biases. Among the 22

[3] The original and final lists of world cities, as well as the literature upon which they are drawn, will be provided upon request.

WCN studies, those pre-dating 2000 have about 70 percent of cities from the United States and Europe while those after 2000 have shifted their focus on cities on other continents, especially Asia and South America. Meanwhile, our approach could inevitably overlook some periphery cities although, as a case study, the focus here is on a list of top world cities. This issue will be further discussed and left for our future work.

Second, a keyword of toponym may be ambiguous. Three types of toponym ambiguation issues may exist in the Google search of city names. First, a city's name may represent more than one city in the world. For example, "Birmingham" is a British city, as well as several other cities with the same name in the United States. Second, a city's name could also represent a state or a country, such as several capital cities (Kuwait City, Mexico City, Luxembourg, and Panama City). Third, a city's name may be a name of famous or popular persons or of other literal meaning. The city "Charlotte" could be a popular person name, while Buffalo could be an animal name. These ambiguation issues are not specific to the webometrics approach, but very common in content analysis and text mining (Smith and Crane 2001; Buscaldi and Rosso 2008; Liu et al. 2014). An effective method to disambiguate geographical names is fulltext geocoding, also known as "geoparsing", or "Geographic Information Retrieval" (GIR) (Jones and Purves 2008; Gelernter and Mushegian 2011). However, such an approach is inappropriate for search engines, since most search engines are keyword-oriented, rather than entity-oriented. We thus create a relatively accurate list of city toponyms as inputs, that is, by adding state, country names or their abbreviations after the city names as the keywords. For example, the keyword of "Birmingham" for searching in Google is "Birmingham" and ("United Kingdom" OR "Britain" OR "UK" OR "England"). A search with more specific keywords may exclude the webpages of ambiguous information and underestimate the importance of top famous cities, such as London, Paris, New York, since it is often unnecessary to mention state or country names when such cities are cited. This might partially flatten out the hierarchy of top-ranking world cities. In contrast, the Google search engine would return more accurate hit counts when the keyword is more concrete (Funahashi and Yamana 2010).

Lastly, the searching algorithms of many search engine companies are not transparent and can return inconsistent results; the hit count of a fixed search may vary over time without explicit explanations (Funahashi and Yamana 2010). To reduce measurement errors, many studies suggest a relatively long-term trace of the searching results (e.g., Bollegala 2008; Liu et al. 2014). Our studies used a similar strategy and collected 28 sets of city network data for one month from 9 May to 6 June, 2016. In total, we conducted 441,000 sets (28*126*125) of searches. Given such a large searching load, we developed a Python-based web-crawling algorithm to automatically search and crawl the hit count data of city pairs. We finally selected the maximum from 28 sets of data as the estimated intercity relations. Also, we focus on the WCN reflected in English webpages only. English ranks first in the ranking of webpage content languages, as estimated by W^{3}Techs.com, and English webpages account for about 56 percent of all webpages.

Analytical Results

City Rankings and Distributions

This study first provides a macroscale analysis of world city ranking and distributions. Figure 2A shows the geographic distribution of world city nodes and top-500 city dyads in the WCN extracted from English webpages. As the size of markers represents the magnitude of degree centrality of cities with edge weights considered, we find that the cities with large degrees are mainly in the Western countries. Among the top-30 cities (Figure 2B), most are Euromerican cities, including 13 in the U.S., 1 in Canada, and 8 in Europe, while there are 5 Asian cities and 3 from Australia and New Zealand. In detail, New York and London rank No. 1 and 2 in degree centrality, implying that they are most referenced with other world cities in English webpages overall. The first non-Euromerican cities in the ranking are Melbourne and Sydney, ranked No. 9 and 10, and the top Asian city is Singapore, with other Asian cities ranking outside the top-20. From the top-100 city-dyad distribution, one can also conceive that most co-references of two cities depict the relations of cities among North America, Europe, or between them; only a limited number of top-100 city-dyads link with cities in the other regions.

[Figure 2 Inserted Here]

In addition to macroregional disparities, the city ranking also communicates a sense of city hierarchy in the WCN (Figure 2B). For example, while the average degree centrality is around 1.9 billion webpages, the degrees of the top-2, New York and London, are 17.8 and 10.3 billion, respectively, over 9 and 5 times the average. Figures 2C-D provide a better understanding on the uneven distributions of the edge weights; 80 percent of weights are concentrated on about 40 percent of city nodes (Figure 2C) and only about 15 percent of city dyads (Figure 2D). Figure 2E shows the cumulative distribution of degree centralities in a log-log plot. If the curve is a straight line, the degree distribution fits a power law, in which the number of nodes of degree k is proportional to $k^{-\beta}$ (Newman 2010). A larger β represents a more hierarchical structure in a WCN. According to Figure 2E, although the curve as a whole does not fit the power law, it can be separated into two linear sections. The first section includes the distribution of low-degree city nodes (about 60 percent of total cities) with degrees below 2 billion (β=1.4), while the second section includes the degree distribution of 40 percent of high-degree city nodes (β=2.9). These degree distributions imply that the degrees are more hierarchically distributed in the top-40 percent cities than the bottom-60 percent cities. According to Chung and Lu (2002), when $\beta \in (2, 3)$, the power-law (or scale-free) network may well contain cohesive cores. Thus, it may be easier to detect dense cores in the subnetwork of top-40 percent cities.

Mesoscale Structure Detection

The best-fit WSBM is identified by changing the number of partitions *K* from 2 to 15, checking how the approximate marginal log-likelihood value changes with *K*, and then detecting the optimal number of partitions which returns the maximum log-likelihood value. In this case study, the optimal number of groups is determined to be 8.

Figure 3A presents the heat-map representation of the pairwise city-based adjacency matrix, where row and columns are indexed by cities sorted by the magnitude of their degree centrality. Given the schematic mesoscale structures that may be revealed (Figure 1), results here suggest that the WCN based on English webpage contents has a mesoscale structure of multiple cores and peripheries, instead of simply a single CP structure or a community structure found in some previous studies (Alderson and Beckfield 2004; Martinus and Sigler 2017; Rozenblat et al. 2017). This WCN is a hybrid of CP and community structures, which implies a mixed configuration of hierarchical and horizontal structures. We can further reduce the city-by-city network to a shrunk block-by-block network to better present the inter-group relations. Figure 3B provides a

3-dimentional block matrix with the vertical axis representing the mean weights of edges within blocks (i.e., $\boldsymbol{\mu}$). If the third group and its connected blocks are removed, one discovers a relatively standard CP structure, in which the average numbers of coreferenced webpages within blocks gradually decrease from about 600 to 1 million from the core to the peripheral blocks. The third group has strong connections with the first two core groups and relatively low linkages with remaining groups, implying that the third group is an independent community from the remaining groups 4-8. Figure 3C depicts a shrunk network with 8 city groups as nodes and the average inter-block relations as edges. To better tease out the hierarchy of the shrunk network, we truncate the top-half edges above the average level of 15.3 million and highlight them in blue. According to the truncated subnetwork, there are 7 degrees for Group 1, 4 for Groups 2 and 4, 3 for Groups 3 and 5, 1 for Groups 6 and 7, and 0 for Group 8. Based on these connection patterns, we define four roles in the WCN: global core (Group 1), macro-region cores (2 and 4), semi-peripheries (3 and 5), and peripheries (6, 7 and 8).

[Figures 3 & 4 Inserted Here]

Figure 4 further visualizes the multi-cores-peripheries (MCP) structure in a network layout and provides the roster of cities in each group. The graph representation shows that, as the only city in Group 1, New York serves as the global core, or α-core, and thus is positioned at the center of the WCN graph. Surrounding New York, there are two macroregional core groups, including 10 cities in a β-Core (i.e., Group 2: London, Paris, Berlin, Rome, Chicago, San Francisco, Toronto, Melbourne, Sydney, New Delhi) and 5 cities in a γ-Core (i.e., Group 4: Singapore, Beijing, Hong Kong, Shanghai, Tokyo). The β-core mainly includes 4 prime European cities and 3 major

U.S. cities, as well as 3 Asia-Pacific cities from the British Commonwealth. In contrast, the γ-Core cities are all major Asian cities. World cities in the three cores are often found in the rosters of top-ranking cities in many WCN studies (Beaverstock, Smith and Taylor 1999; Alderson and Beckfield 2004; Taylor and Derudder 2015). In contrast with ranking studies, however, our clustering findings use the link patterns between groups to define the roles of city groups. These cores have the same link patterns that, within the cores, cities are highly connected while the relations between the cores and peripheries are fairly strong. Although their connection patterns look similar, the WSBM can still distinguish the three cores. This is mainly ascribed to different scales of connections. As shown in Figure 3B, the α-core has the highest mean weight linking with each of the other groups; the β-core has significantly lower mean weight than the α-core but larger than the γ-core. This configuration of multiple cores is understated in the existing WCN literature.

While many U.S. cities rank high overall (Figure 1), most are not identified as the core cities in the WCN here since they have different relational patterns compared to the core groups. This result implies that the clusters based on rankings are relatively preliminary and may be misleading. Focusing on two semi-periphery groups (Groups 3 and 5), both have a relatively high average level of connection with the three cores (over 4 times the average edge weight) but rather weak linkage to each other and to other peripheral groups (ranging from 7 to 47 percent of the average weight). Figure 4 clearly shows two distinctive communities of semi-peripheries that are well connected to the cores, but those within-community links are much denser than those between the two communities. In detail, the semi-periphery#1 group (Group 3) is dominated by U.S. cities (15, or 88 percent), supplemented by two European cities (12 percent); this group is particularly well connected with the two Euromerican cores (α-core and β-core). On the other hand, U.S. cities are conspicuously absent from the semi-periphery#2 group (Group 5), which is formed of 23 cities from diverse regions, including Europe (57 percent), South America (17 percent), Asia (17 percent), Oceania (4 percent) and Africa (4 percent); this group is well connected with all three cores.

The three peripheries have a role in the WCN that is somewhat similar to the semi-peripheries, but they have relatively weak connections with both the cores and other peripheries. They account for about 56 percent of total cities in this study. This finding appears consistent with the macroscale analysis reported in the previous section that the bottom 60 percent cities may be less hierarchical than others. Figure 4 conveys a better sense of the similarity and difference between the three peripheries. They are all circularly distributed around the cores and semi-peripheries, with the more peripheral groups farther away from the cores than the less peripheral groups. None of the peripheral groups are geographically cohesive, as all the groups include cities from different continents.

Comparative Mesoscale Analysis

Geographers have articulated diverse imaginaries of the mesoscale structures of the world, including a CP hierarchy (Friedmann 1986), a flat world (Friedman 2005), a spiky world (Florida 2008), and a fuzzy world regionalization (Taylor et al. 2013). While these geographical imaginaries are all intuitively appealing, their validation remains work in progress and their comparative examination has been difficult under the existing analytical perspective. The new approach raised here particularly encourages to investigate whether, and how much, a WCN looks like a CP hierarchy or some other structures. An important advantage of WSBM analysis is allowing researchers for statistical or probability-based comparison of empirically derived mesoscale structures with some presumed structures. Following Eq. (2), the log-likelihood value of the optimal partition, representing the data fit of the best possible model, with the MCP structure (Figure 4) is calculated as $LL_{\mathrm{op}} = -1.275 \times 10^5$. For comparison, we can define a null model without any partition (i.e., a flat world, or *K*=1), representing the fit of the worst possible model in statistics, and compute its log-likelihood score as $LL_0 = -1.439 \times 10^5$. Comparing the former to the latter, we find that the WCN data in this study are significantly more likely to fit into a MCP structure, as shown in Figure 4, than the null model that assumes a flat world.

In fact, the observed WCN possibly has other mesoscale structures. It is thus interesting to compare different mesoscale structures and identity to what extent a specific partition or mesoscale structure differs from the best partition or the null model. Accordingly, we define the following incremental fit index (IFI), given the log-likelihood of a partition of mesoscale structure *x*:

$$IFI_x = \frac{LL_x - LL_0}{LL_{\mathrm{op}} - LL_0} \tag{4}$$

The IFI is analogous to R^2 in regression models. A value near zero indicates that the presumed mesoscale structure *x* is close to the null model without partitions such that a partition *x* is statistically unsound, while a value close to 1 indicates that *x* is close to the best partition of mesoscale structure.

Figure 5 shows the heat maps of the city-by-city matrices for two partitioning hypotheses. One partitions cities into 5 groups based on the ranking of weighted degree centralities (i.e., the ranking model in Figure 5A). While there are countless ways of partitioning by city ranks, we simply set the boundaries of bins based on the breaking points with a sharp change of intermediate slope for coarse comparison. The other one classifies cities into 6 groups by continent (i.e., the regionalization model in Figure 5B). This regionalization (or territorialist-geography) approach often assumes or explains a WCN of relatively horizontal community structure such that cities within a specific geographic boundary (e.g., a continent) are more similar or strongly connected with each other. The IFI of the ranking model is 83.8 percent, showing that the WCN in this study is more likely to exhibit a MCP hierarchy than a city-ranking hierarchy. A partition by city ranks may not demonstrate the best-fit mesoscale structures in the WCN, although the ranking partition appears more similar to the MCP structure than the flat-world model.

In contrast, the regionalization model only has an IFI of 28.4 percent, implying that the WCN reflected in webpage contents has no significant territorialist properties. Many world cities play a role in the WCN outside their national and even continental (or macro-regional) realm, and those cities embedded in a similar network position may come from different territories. This finding contradicts some clustering studies (e.g., Derudder et al. 2003; Liu et al. 2014; Neal 2014; Martinus and Sigler 2017; Rozenblat et al. 2017), which often found a regional pattern and tendency of world city clusters. This difference may result from different relational data sources: regional patterns are mainly found in a WCN based on the APS corporate organization. Compared to this, the geographical connections among cities transpires relatively less in the broad-based WCN in this study. Also, traditional clustering methods do not provide unambiguous criteria for partitions and thus would easily result in a bias towards "methodological territorialism".

Discussion and Conclusions

This study develops a new conceptual framework for WCN studies that focuses on mesoscale structures in WCNs and recognizes the plurality of such structures, namely hierarchical, territorialist, and hybrid structures. As such, it is a shift from conventional approaches that have focused either on local or on global structures. Yet, this should not be understood as advocacy for the abandonment of the traditional microscale and macroscale analyses, but rather for the incorporation of mesoscale analysis to more comprehensively understand world cities, WCN structures, and globalization. This is accomplished by studying latent mesoscale structures (e.g., CP hierarchies, communities, or hybrids) in relational networks. Methodologically, this is the first study to develop a probability-based WSBM for comparative mesoscale analysis of WCNs. We show the importance and advantages of the new analytical framework in a case study in which the WCN is configured by the coreference frequency of each city pair on Internet webpages. The WCN data was collected by a webometrics approach, including Google-Search-based web crawling and content analysis and disambiguation.

Modeling results suggest that the WCN reflected in webpage contents has a multi-cores-peripheries (MCP) structure mixed with communities. Through comparative analysis, we estimate the probability of different mesoscale structures and conclude that the globalization of world cities reflected in a big dataset of English webpages is less likely to be a flat-world process as imaged by Friedman (2005), or a purely ranking hierarchy as presumed by Friedmann (1986) or traditional "territorialist" regional geographies (Derudder et al. 2003). These findings partially align with Taylor et al.'s (2013) "new regional geographies" that deny a homogenous world and describe an overlapping world regionalization, in which the overlapped cities are in effects core cities.

The main contribution of this article lies in the new analytical perspective and methodology to the WCN literature, enabling us to uncover the significance of the MCP hybrid structure, for which further exploration and comparison are in order in the future. Many advantages of the new analytical framework have been found. First, it can avoid the trap of "*methodological determinism*". It neither postulates the existence of hierarchies as some CP-detection methods do (e.g., Alderson and Beckfield 2004) or communities as the community-detection algorithms do (e.g., Martinus and Sigler 2017; Rozenblat et al. 2017), nor simply ignores the likelihood of a hierarchical structure (e.g., Coe et al. 2010). The identification of hierarchies and other mesoscale structures requires to look at grouping properties of city nodes and the connection patterns within and between groups, rather than focusing on individual city's nodes or dyads as most existing WCN studies do. Second, the new approach can open the "black box" of networks by disentangling alternative mesoscale structures to hierarchies situated in WCNs, which is vaguely examined by the dualism tradition. This also provides a new approach to comparing WCNs of different intercity relations across temporal and/or spatial contexts. Third, it can avoid the trap of "methodological territorialism" (Brenner 1999) since the regional or territorial scopes of city clusters are not predetermined but derived by the pattern of power relations between cities that sort them into regions or territories. Lastly, since there is a long tradition asking for the integration of network analysis into WCN studies (Smith and Timberlake 1995), the new analytical framework helps to bridge WCN studies and the fast-growing field of network science.

There is room for improving this research, especially the case study. First, while the co-referencing data used here is capable for measuring an aggregation of various linkages between cities, they may fail to tell the details of what these linkages are. For a better interpretation of the co-referencing networks, our future work will classify the intercity linkages based on different webpage contents. Recent examples include studies by Salvini and Fabrikant (2015) and Hu et al. (2017) on semantic relatedness between cities, in which the intercity relations of economy/technology, politics, cultural, and human-interests relations are identified using Wikipedia document and news articles data. The classification can also help us to better compare the document approach with the corporate and infrastructure approaches. Second, since the case study only considers English webpages, the analytical results may mainly reflect the globalization imaginaries of English-speaking countries and underestimate the positions and roles of cities in non-English countries, like most Asian cities (Devriendt et al. 2011; Salvini and Fabrikant 2015). This article in effect provides a promising method to compare the mesoscale structures of the WCNs reflected in webpages by different languages in the future. Also, it appears that the mesoscale analysis can mitigate the selection bias, compared to the ranking-based analysis. For instance, it has been documented that some major Asian cities play increasingly important roles in the WCN and often rank in the top 10 in some recent GaWC studies based on APS corporate linkages (e.g., Taylor and Derudder 2015). Given the known language bias, the centrality ranking method is ill suited to recognize the “true” roles of these Asian cities --they rank out of the top10 in this study (Figure 2), whereas the mesoscale analysis successfully identifies 5 Asian cities serving as one of three cores (γ-Core). Third, as discussed before, the detected mesoscale structures may be partially pre-determined by the selection of world cities and the network data for WCN studies; thus, further studies should broaden the number of world cities under consideration, while accounting for disparities across world regions. Notwithstanding these limitations, the mesoscale analytical perspective and methodological framework proposed here add substantive value to the WCN debate and can be also applied to conventional WCN data, such as the APS, infrastructure, and industrial corporation networks.

Acknowledgments

We thank the reviewers and the editor Mei-Po Kwan for their insightful comments, as well as Christopher Aicher for providing Matlab codes to run the WSBM methods.

References

Aicher, C., A.Z. Jacobs, and A. Clauset. 2014. Learning latent block structure in weighted networks. *Journal of Complex Networks* 3(2):221-248. doi: 10.1093/comnet/cnu026.

Alderson, A.S., and J. Beckfield. 2004. Power and Position in the World City System 1. *American Journal of Sociology* 109(4):811-851. doi: 10.1086/378930.

Almind, T.C., and P. Ingwersen. 1997. Informetric analyses on the world wide web: methodological approaches to 'webometrics'. *Journal of Documentation* 53(4):404-426. doi: 10.1108/eum0000000007205.

Berry, B.J. 1961. City size distributions and economic development. *Economic Development and Cultural Change* 9 (4, Part 1): 573-588. doi: 10.1086/449923.

Beaverstock, J.V., R.G. Smith, and P.J. Taylor. 1999. A roster of world cities. *Cities* 16(6):445-458. doi: 10.1016/s0264-2751(99)00042-6.

Beaverstock, J.V., R.G. Smith, and P.J. Taylor. 2000. World□city network: A new metageography? *Annals of the Association of American Geographers* 90(1):123-134. doi: 10.1111/0004-5608.00188.

Beaverstock, J.V., R.G. Smith, P.J. Taylor, D.R.F. Walker, and H. Lorimer. 2000. Globalization and world cities: some measurement methodologies. *Applied Geography* 20(1):43-63. doi: 10.1016/s0143-6228(99)00016-8.

Beck, U. 2007. The cosmopolitan condition: Why methodological nationalism fails. *Theory, Culture & Society* 24(7-8):286-290. doi: 10.1177/02632764070240072505.

Beckfield, J., and A.S. Alderson. 2004. Power and position in the world city system. *American Journal of Sociology* 109(4):811-851. doi: 10.1086/378930.

——— 2006. Reply: Whither the Parallel Paths? The Future of Scholarship on the World City

System 1: See the comment associated with this reply. *American Journal of Sociology* 112(3):895-904. doi: 10.1086/510992.

Bollegala, D., T. Honma, Y. Matsuo, and M. Ishizuka. 2008. Mining for personal name aliases on the web. Paper presented at the 17th International Conference on World Wide Web, Beijing, China, April 21-25.

Borgatti, S.P., and M.G. Everett. 2000. Models of core/periphery structures. *Social networks 21*(4):375-395. doi: 10.1016/s0378-8733(99)00019-2.

Brenner, N. 1999. Beyond state-centrism? Space, territoriality, and geographical scale in globalization studies. *Theory and Society* 28(1):39-78. doi: 10.1023/A:1006996806674.

Brown, E., B. Derudder, C. Parnreiter, W. Pelupessy, P.J. Taylor, and F. Witlox.2010. World City Networks and Global Commodity Chains: towards a world□systems' integration. *Global Networks* 10(1):12-34. doi: 10.1111/j.1471-0374.2010.00272.x.

Buscaldi, D., and P. Rosso. 2008. A conceptual density□based approach for the disambiguation of toponyms. *International Journal of Geographical Information Science* 22(3):301-313. doi: 10.1080/13658810701626251.

Coe, N.M., P. Dicken, M. Hess, and H.W.C. Yeung. 2010. Making connections: global production networks and world city networks. *Global Networks* 10(1):138-149. doi: 10.1111/j.1471-0374.2010.00278.x.

Csermely, P., A. London, L.Y. Wu, and B. Uzzi. 2013. Structure and dynamics of core/periphery networks. *Journal of Complex Networks* 1(2):93-123. doi: 10.1093/comnet/cnt016.

Derudder, B. 2008. Mapping global urban networks: a decade of empirical world cities research. *Geography Compass* 2(2):559-574. doi: 10.1111/j.1749-8198.2008.00102.x.

Derudder, B., and F. Witlox. 2005. An appraisal of the use of airline data in assessing the world city network: a research note on data. *Urban Studies* 42(13):2371-2388. doi: 10.1080/00420980500379503.

Derudder, B., and P.J. Taylor. 2017. Central flow theory: comparative connectivities in the world-city network. *Regional Studies* 1-14. doi: 10.1080/00343404.2017.1330538.

Derudder, B., P.J. Taylor, F. Witlox, and G. Catalano. 2003. Hierarchical tendencies and regional patterns in the world city network: a global urban analysis of 234 cities. *Regional Studies* 37(9):875-886. doi: 10.1080/0034340032000143887.

Devriendt, L., A. Boulton, S. Brunn, B. Derudder, and F. Witlox. 2011. Searching for cyberspace: the position of major cities in the information age. *Journal of Urban Technology* 18(1):73-92. doi: 10.1080/10630732.2011.578410.

Faust, K., and S. Wasserman. 1992. Blockmodels: Interpretation and evaluation. *Social Networks* 14(1-2):5-61. doi: 10.1016/0378-8733(92)90013-w.

Funahashi, T., and H. Yamana. 2010. Reliability verification of search engines" hit counts: How to select a reliable hit count for a query. *Current Trends in Web Engineering* 114-125. doi: 10.1007/978-3-642-16985-4_11.

Florida, R. 2008. *Who's your city?* New York, NY: Basic Books.

Fortunato, S. 2010. Community detection in graphs. *Physics reports 486*(3):75-174. doi: 10.1016/j.physrep.2009.11.002.

Fortunato, S., and D. Hric. 2016. Community detection in networks: A user guide. *Physics Reports* 659:1-44. doi: 10.1016/j.physrep.2016.09.002.

Friedman, T.L. 2005. *The World Is Flat: A Brief History Of The Twenty-First Century*. New York, NY: Macmillan.

Friedmann, J. 1986. The world city hypothesis. *Development and Change 17*(1):69-83. doi:

10.1017/cbo9780511522192.019.

———. 1995. Where we stand: A decade of world city research. In *World Cities in A World-System,* ed. P. L. Knox and P. J. Taylor, 21–47. Cambridge, U.K.: Cambridge University Press.

Gelernter, J., and N. Mushegian. 2011. Geo-parsing messages from microtext. *Transactions in GIS* 15(6):753-773. doi: 10.1111/j.1467-9671.2011.01294.x.

Godfrey, B.J., and Y. Zhou. 1999. Ranking world cities: multinational corporations and the global urban hierarchy. *Urban Geography* 20(3):268-281. doi: 10.2747/0272-3638.20.3.268.

Hall, P. 1966. *The World Cities*. London, U.K.: Heinemann.

Holland, P.W., K.B. Laskey, and S. Leinhardt. 1983. Stochastic blockmodels: First steps. *Social Networks* 5(2):109-137. doi: 10.1016/0378-8733(83)90021-7.

Hu, Y., X. Ye, and S-L. Shaw. 2017. Extracting and analyzing semantic relatedness between cities using news articles. *International Journal of Geographical Information Science* 31(12):427-2451. doi: 10.1080/13658816.2017.1367797.

Huntington, S.P. 1997. *The clash of civilizations and the remaking of world order*. New York, NY: Penguin Books India.

Hoyler, M., O. Mould, and P. Taylor. 2018. Global Buzz – Measuring Short-Term Changes Among Cities in Globalization. http://www.lboro.ac.uk/gawc/projects/projec88.html (accessed May 8, 2018).

Hymer, S. 1972. The Multinational Corporation and the Law of Uneven Development. In *Economics and World Order from the 1970s to the 1990s*, ed. J. Bhagwati, 113 – 40. London, U.K.: Collier-Macmillan.

Jones, C.B., and R.S. Purves. 2008. Geographical information retrieval. *International Journal of Geographical Information Science* 22(3):219-228. doi: 10.1080/13658810701626343.

Liu, X., B. Derudder, and P. Taylor. 2014. Mapping the evolution of hierarchical and regional tendencies in the world city network, 2000–2010. *Computers, Environment and Urban Systems* 43:51-66. doi: 10.1016/j.compenvurbsys.2013.10.004.

Liu, Y., F. Wang, C. Kang, Y. Gao, and Y. Lu. 2013. Analyzing relatedness by toponym co-occurrences on web pages. *Transactions in GIS* 18(1):89-107. doi: 10.1111/tgis.12023.

Lloyd, P., M.C. Mahutga, and J. De Leeuw. 2009. Looking back and forging ahead: Thirty years of social network research on the world-system. *Journal of World-Systems Research* 15(1):48-85. doi: 10.5195/jwsr.2009.335.

Jeub, L.G., P. Balachandran, M.A. Porter, P.J. Mucha, and M.W. Mahoney. 2015. Think locally, act locally: Detection of small, medium-sized, and large communities in large networks. *Physical Review E* 91(1): 012821. doi: 10.1103/physreve.91.012821.

Karrer, B., and M.E. Newman. 2011. Stochastic blockmodels and community structure in networks. *Physical Review E* 83(1):016107. doi: 10.1103/physreve.83.016107.

Kick, E.L., L.A. McKinney, S. McDonald, and A. Jorgenson. 2011. A multiple-network analysis of the World System of Nations, 1995-1999.In *Sage handbook of social network analysis*, ed. J. Scott, and P.J. Carrington, 311-328. Los Angeles, California: SAGE Publications Ltd.

Mans, U. 2014. Revisiting city connectivity. *Journal of Economic Geography* 14(1): 155–177. doi: 10.1093/jeg/lbs054.

Martinus, K., T.J. Sigler, G. Searle, and M. Tonts. 2015. Strategic globalizing centers and sub-network geometries: A social network analysis of multi-scalar energy networks.

Geoforum 64:78-89. doi: 10.1016/j.geoforum.2015.06.006.

Martinus, K., and M. Tonts. 2015. Powering the world city system: energy industry networks and interurban connectivity. *Environment & Planning A: Economy and Space* 47(7):1502-1520. doi: 10.1177/0308518x15599124.

Martinus, K., and T.J. Sigler. 2017. Global city clusters: theorizing spatial and non-spatial proximity in inter-urban firm networks. *Regional Studies* 1-13. doi.org/10.1080/00343404.2017.1314457

Mahutga, M.C., X. Ma, D.A. Smith, and M. Timberlake. 2010. Economic globalisation and the structure of the world city system: the case of airline passenger data. *Urban Studies* 47(9):1925-1947. doi: 10.1177/0042098010372684.

Neal, Z. 2012. *The connected city: How Networks are Shaping the Modern Metropolis*. New York, NY: Routledge.

———. 2012. Structural determinism in the interlocking world city network. *Geographical Analysis* 44(2):162-170. doi: 10.1111/j.1538-4632.2012.00843.x.

———. 2014. The devil is in the details: differences in air traffic networks by scale, species, and season. *Social Networks* 38(3):63-73. doi: 10.1016/j.socnet.2014.03.003.

———. 2015. Does world city network research need eigenvectors? *Urban Studies* 50(8):1648-1659. doi: 10.1177/0042098013477702.

Newman, M.E., and M. Girvan. 2004. Finding and evaluating community structure in networks. *Physical Review E* 69(2): 026113. doi: 10.1103/physreve.69.026113.

Pred, A.R. 1977. *City systems in advanced economies: Past growth, present processes, and future development options*. New York, NY: John Wiley & Sons.

———. 1980. *Urban growth and city systems in the United States, 1840–1860.* London, U.K.:

Hutchinson.

Rombach, M.P., M.A. Porter, J.H. Fowler, and P.J. Mucha. 2014. Core-periphery structure in networks. *SIAM Journal on Applied Mathematics* 74(1):167-190. doi: 10.1137/120881683.

Robinson, J. 2002. Global and world cities: a view from off the map. *International Journal of Urban and Regional Research* 26(3):531-554. doi: 10.1111/1468-2427.00397.

Rossi, E.C., J.V. Beaverstock, and P.J. Taylor. 2007. Transaction links through cities: 'decision cities' and 'service cities' in outsourcing by leading Brazilian firms. *Geoforum* 38(4):628-642. doi: 10.1016/j.geoforum.2006.11.005.

Rozenblat, C., F. Zaidi, and A. Bellwald. 2017. The multipolar regionalization of cities in multinational firms' networks. *Global Networks* 17(2):171–194. doi: 10.1111/glob.12130.

Sassen, S. 1991. *The Global City: New York, London, Tokyo*. Princeton, NJ: Princeton.

Salvini, M.M., and S.I. Fabrikant. 2015. Spatialization of user-generated content to uncover the multirelational world city network. *Environment & Planning B Planning & Design* 43(1):228-248. doi: 10.1177/0265813515603868.

Smith, D.A., and G. Crane. 2001. Disambiguating geographic names in a historical digital library. In *International Conference on Theory and Practice of Digital Libraries*, ed. 127-136. Springer, Germany: Berlin Heidelberg.

Smith, D.A., and M. Timberlake. 1995. Conceptualising and mapping the structure of the world systems city system. *Urban Studies* 32(2):287–302. doi: 10.1080/00420989550013086.

Snyder, D., and E.L.K ick. 1979. Structural position in the world system and economic growth, 1955-1970: A multiple-network analysis of transnational interactions. *American journal of Sociology* 84(5):1096-1126. doi: 10.1086/226902.

Taylor, P. J. 1997. Hierarchical tendencies amongst world cities: a global research proposal.

Cities 14(6):323–332. doi: 10.1016/s0264-2751(97)00023-1.

———. 2004. *World city network: a global urban analysis (1st Edition)*. Abingdon, U.K.: Routledge.

———. 2005. New political geographies: Global civil society and global governance through world city networks. *Political Geography 24*(6):703-730. doi: 10.1016/j.polgeo.2005.01.009.

———. 2006. Comment: Parallel Paths to Understanding Global Intercity Relations 1: See the reply associated with this comment. *American Journal of Sociology* 112(3):881-894. doi: 10.1086/510991.

———. 2009. Urban economics in thrall to Christaller: a misguided search for city hierarchies in external urban relations. *Environment and Planning A* 41(11):2550-2555. doi: 10.1068/a42235.

Taylor, P.J., and B. Derudder. 2015. *World city network: a global urban analysis (2st version)*. Abingdon, U.K.: Routledge.

Taylor, P.J., B. Derudder, M. Hoyler, and P. Ni. 2013. New regional geographies of the world as practised by leading advanced producer service firms in 2010. *Transactions of the Institute of British Geographers* 38(3):497-511. doi: 10.1111/j.1475-5661.2012.00545.x.

Taylor, P.J., and D.R.F. Walker. 2001. World cities: a first multivariate analysis of their service complexes. *Urban Studies* 38(1):23-47. doi: 10.1080/00420980125400.

Thelwall, M. 2009. Introduction to webometrics: Quantitative web research for the social sciences. *Synthesis Lectures on Information Concepts, retrieval, and services 1*(1):1-116. doi: 10.2200/s00176ed1v01y200903icr004.

Therborn, G. 2011. End of a paradigm: the current crisis and the idea of stateless cities.

Environment and Planning A 43(2):272-285. doi: 10.1068/a42328.

Tunç, B., and R. Verma. 2015. Unifying inference of meso-scale structures in networks. *Plos One* 10(11):e0143133. doi: 10.1371/journal.pone.0143133.

Vaughan, L., and M. Thelwall. 2004. Search engine coverage bias: evidence and possible causes. *Information Processing & Management* 40(4):693-707. doi: 10.1016/s0306-4573(03)00063-3.

Wall, R.S., and G.A. Van der Knaap. 2011. Sectoral differentiation and network structure within contemporary worldwide corporate networks. *Economic Geography 87*(3):267-308. doi: 10.1111/j.1944-8287.2011.01122.x.

Wallerstein, I. 1974. The rise and future demise of the world capitalist system: concepts for comparative analysis. *Comparative Studies in Society and History* 16(04):387-415. doi: 10.1017/s0010417500007520.

Wasserman, S., and K. Faust. 1994. *Social network analysis: Methods and applications (Vol. 8)*. Cambridge, U.K.: Cambridge University Press.

White, H.C., S.A. Boorman, and R.L. Breiger. 1976. Social structure from multiple networks. I. Blockmodels of roles and positions. *American Journal of Sociology* 81(4):730-780. doi: 10.1086/226141.

Wimmer, A., and N. Schiller. 2002. Methodological nationalism and beyond: nation–state building, migration and the social sciences. *Global Networks 2*(4):301-334. doi: 10.1111/1471-0374.00043.

Author information

Wenjia Zhang

WENJIA ZHANG is an Assistant Professor in the School of Urban Planning & Design at Peking University Shenzhen Graduate School, Shenzhen, Guangdong 518055, China. Email: zhangwj@pkusz.edu.cn. His research interests include network complexity in urban and geography studies, network and spatial economics, land use and transportation planning, as well as big data and machine learning approaches.

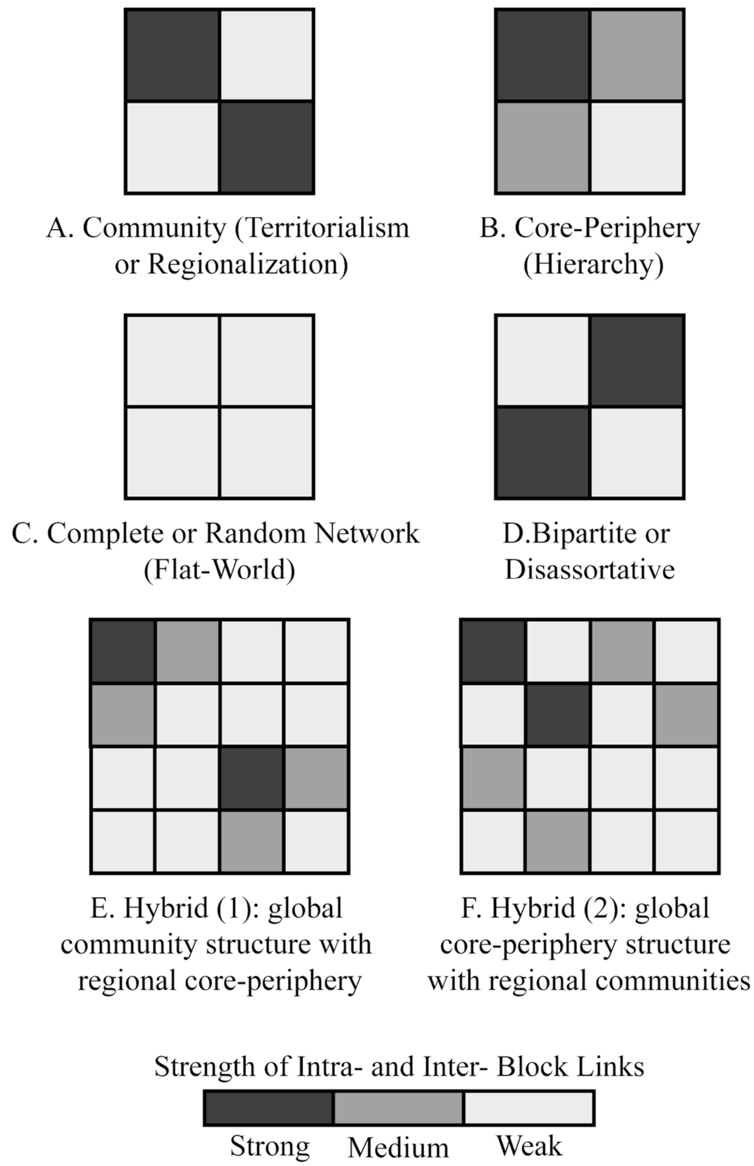
Jean-Claude Thill

JEAN-CLAUDE THILL is Knight Distinguished Professor of Public Policy in the Department of Geography & Earth Sciences at the University of North Carolina at Charlotte, Charlotte, NC 28221. Email: jean-claude.thill@uncc.edu. His research interests focus on relational spaces constructed at various spatial and temporal resolutions, as they express the lived experiences of people, businesses, and other human agencies, through commerce, daily commuting, territorial control by non-state agents, and many other social and behavioral modalities.

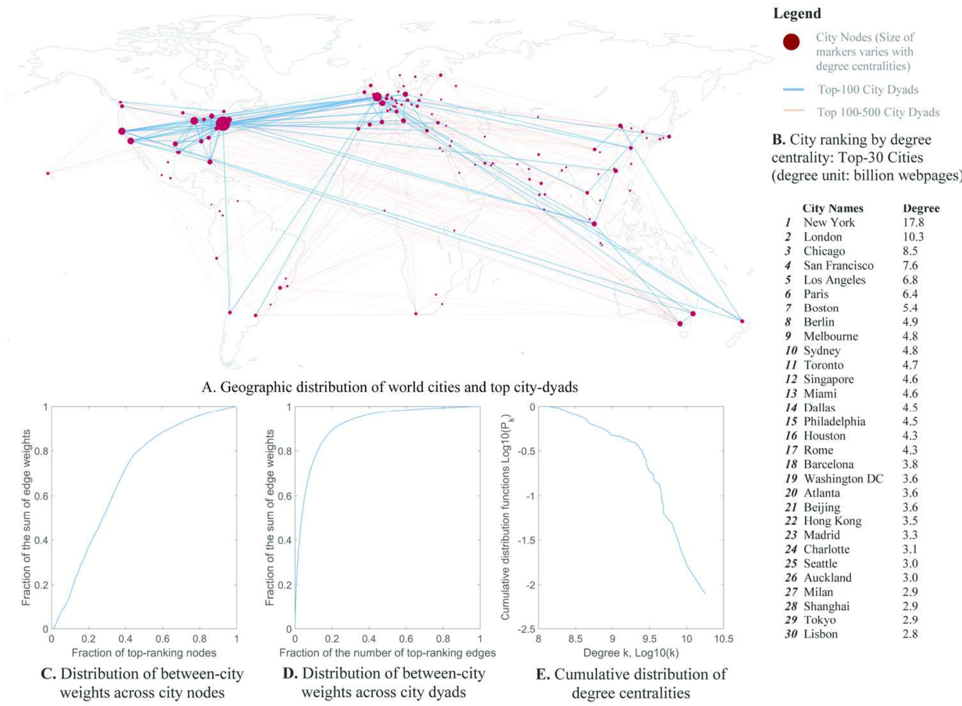
List of Figures

Notes: Colors represent groups. The marker sizes of city nodes and labels represent the level of weighted degree centrality. The width of between-city edges represents the strength of city dyads. The figure is created in a Fruchterman-Reingold layout in Gephi 0.9.1 (https://gephi.org).

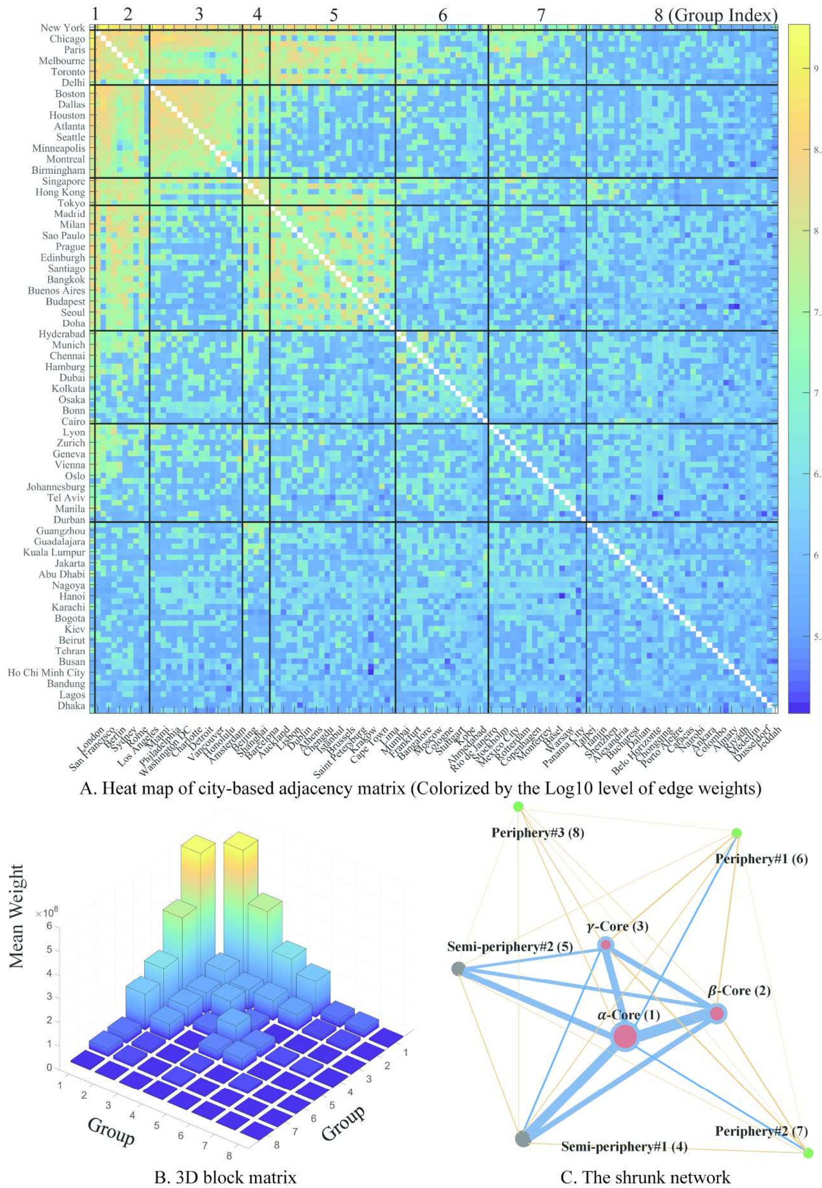
Notes: The log-likelihood values (LL) estimated by Eq. (2) of the optimal partitions derived by the WSBM analysis (Figure 3a) and the null model without any partitions (K=1) are $LL_{\mathrm{op}} =$ -127528.91 and $LL_0 =$ -143949.42. The Incremental Fit Index (IFI) equals $(LL_{\mathrm{presumed}} - LL_0)/(LL_{\mathrm{op}} - LL_0)$.



112x165mm (300 x 300 DPI)



165x119mm (300 x 300 DPI)

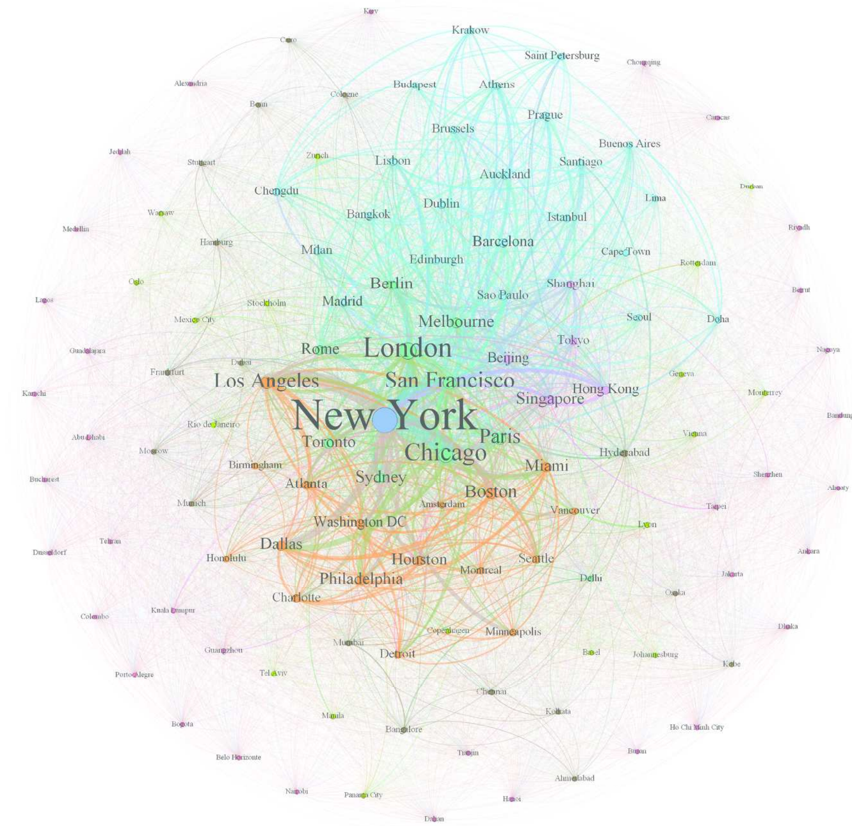


A. Heat map of city-based adjacency matrix (Colorized by the Log10 level of edge weights)

B. 3D block matrix

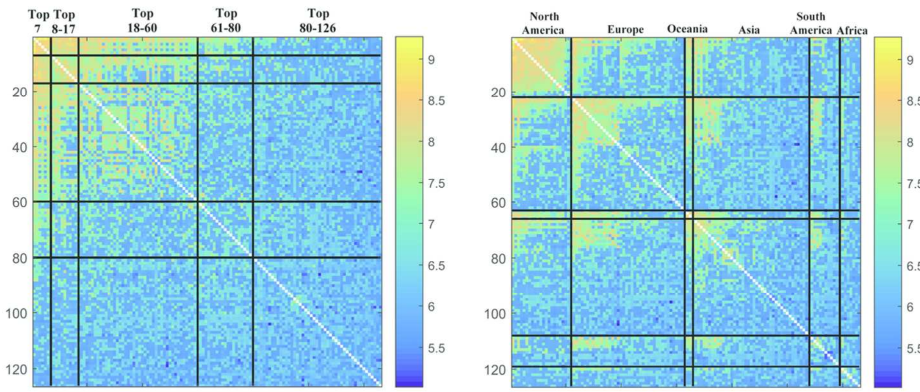
C. The shrunk network

228x293mm (300 x 300 DPI)



214x259mm (300 x 300 DPI)

Group (City No.)	Cities	Roles
1 (1)	New York	α-Core
2 (10)	London, Paris, Berlin, Rome, Chicago, San Francisco, Toronto, Melbourne, Sydney, New Delhi	β-Core
3 (17)	Los Angeles, Boston, Miami, Dallas, Philadelphia, Houston, Washington DC, Atlanta, Charlotte, Seattle, Detroit, Minneapolis, Honolulu, Vancouver, Montreal, Birmingham, Amsterdam	Semi-periphery #1
4 (5)	Singapore, Beijing, Hong Kong, Shanghai, Tokyo	γ-Core
5 (23)	Barcelona, Madrid, Milan, Lisbon, Dublin, Prague, Athens, Brussels, Edinburgh, Istanbul, Saint Petersburg, Budapest, Krakow; Sao Paulo, Buenos Aires, Santiago, Lima; Seoul, Bangkok, Chengdu, Doha; Auckland; Cape Town	Semi-periphery #2
6 (17)	Munich, Frankfurt, Hamburg, Stuttgart, Cologne, Bonn, Moscow; Mumbai, Chennai, Bangalore, Dubai, Kolkata, Hyderabad, Ahmedabad, Osaka, Kobe; Cairo	Periphery#1
7 (18)	Lyon, Stockholm, Zurich, Geneva, Rotterdam, Vienna, Copenhagen, Oslo, Basel, Tel Aviv, Warsaw; Rio de Janeiro, Mexico City, Monterrey, Panama City; Johannesburg, Durban; Manila	Periphery#2
8 (35)	Taipei, Guangzhou, Tianjin, Shenzhen, Chongqing, Dalian, Kuala Lumpur, Jakarta, Abu Dhabi, Nagoya, Hanoi, Karachi, Beirut, Tehran, Colombo, Busan, Almaty, Ho Chi Minh City, Riyadh, Bandung, Dhaka, Jeddah; Guadalajara, Belo Horizonte, Porto Alegre, Bogota, Caracas, Medellin; Bucharest, Kiev, Dusseldorf, Ankara; Alexandria, Nairobi, Lagos	Periphery#3



A. The ranking model: five partitions based on the city ranking of weighted degree centralities (Breaking points for grouping are based on the points with a sharp change of intermediate slope. LL=-130189.13. IFI=83.8%)

B. The regionalization model: six partitions based on the regionalization by Continent.(LL=-139282.29. IFI= 28.42%)

119x59mm (300 x 300 DPI)