MAPPING GLOBAL URBAN INTERCONNECTEDNESS THROUGH GOOGLE TREND: A NEW WORLD CITY NETWORK

Zhiqiang WU [1], Jing HAN[2], Wei LIU[3], Qiming YE[4]

[1]College of Architecture and Urban Planning, Tongji University, Shanghai, China, prof.wus@gmail.com
[2]Shanghai Tongji Urban Planning and Design Institute, hjbm333@163.com
[3]College of Architecture and Urban Planning, Tongji University, Shanghai, China, nrliuwei@126.com
[4]College of Architecture and Urban Planning, Tongji University, Shanghai, China, qiming_ye@qq.com

Keywords: Global Urban Network, Information Flow, Google Trends

Abstract:
Network of City under the impact of globalization and information technology is an active direction for the last decade of Urban and Regional Studies. Based upon the search interest statistics provided by Google Trends, this article gets people s online attention *data among cities in 2012, maps the main linkage among world cities, and ranks cities according to their international connectedness. Main results confirm that 1) Europe has the largest number of higher-ranking cities in the new global urban network, the next is North America, and the least is South America as well as Africa; 2) Southeast Asian cities except Singapore and Hong Kong all perform poorly under the new criteria; 3) Two types of cities, info-flow-in and info-flow-out, nearly distribute evenly across the main continents; 4) Connection within and between Europe and North America is of the most density and frequency among other intercontinental links. Overall, significant differences exist between the new networks based upon inter-city concern on the Internet and the traditional urban system based on the producer service firms. Interconnectedness of cities to a large extent is influenced by the real geographic location and distance. The research does a beneficial attempt on mining information flow data and provides new insights into city network studies.*

1. Introduction

World Cities is always a key issue in the field of urban and regional studies. A series of changes have taken place over time in the research methods and ideas. Generally, the mainstream researches can be divided into three stages: stage of world cities with attributes but without relations, stage of world cities in network society, new stage of world city network in global environment (Yang, 2011).

In the first stage (1960s-1990s), most research literatures are about the comparison of the city's comprehensive index (e.g. the concentration of headquarters of multinational companies, controlling rate of financial assets, etc.). The cities ranking in front according to these attributes are defined as world cities. In this period, World Cities study has obtained a great deal of information about urban nodes, but mainly focused on characteristics of cities in and by themselves, n among world cities. (Hall, 1966; Friedmann s world city hypotheses , 1986; Sassen's triad of world cities, 1991; 1994; 1996)

In the second stage (1990s), the development of information technology had a more and more profound impact on the global urban system organized by world cities. Cities were involved into global urban system through the information network, and their operation is increasingly affected by globalization. Castells (1996) proposed the famous theory of "space of flows", revolutionarily distinguished from the traditional "space of place", and defined the concept of nodes and hubs. Manuel Castells' *Global Spaces of Flows* describes the world city from the spatial perspective of global flows as the nodes and hubs with the "most direct influence" on a global scale. He believes that global city s concept is not a static place, but a process of connecting. Global city in world network is the centre in the connecting process of advanced services production and consumption. Castells has built a research framework based on relations of world cities, but the empirical data is lacking at that time. So it s necessary to strengthen the collecting of the relation dat system.

In the third stage (Since 2000), academic studies of world cities has changed from the traditional approaches to a new empirical phase. In this period, there are two main types in the analysis of relational data between cities: two-mode dataset and one-mode dataset:

For the two-mode dataset, researches mainly analyse producer services from the diversity of urban serving function and make use of hierarchical information flow between multinationals instead of the real connection between cities. Liu and Derudder (2013) systematically summarized three existing calculating approaches of urban network (the ownership linkage model, the interlocking network model, and the two-mode city-by-firm network model). By interlocking network model method, Taylor (2001) measured the strength of the force in global urban network through a large number of case studies, pointed out the key elements in the formation process of global urban system, and built the bridge from enterprise s linkage to cities linkage, which has become the theo
foundation stone of global urban network studies by corporation. Nevertheless, the relational data of Taylor s interlocking network model is unable to express directional link betwe
different connecting characteristics resulted from enterprise s various industries, sizes and strategies might bring biases. To solve this problem, the two-mode city-by-firm network model is applied to evaluate firms and cities positions in corporate networks simultaneously. But t
approach and the associated computational burden deters urban scholars from using it.

For the one-mode dataset, researches reflect urban network from some measurable aspects of economic activity between cities, such as the infrastructure s connections, e.g. air flow (Derudder and Witlox, 2008; Smith and Timberlake, 1995; 2001), telecommunications streaming, internet connections etc. (Malecki, 2002; Tranos and Gillespie, 2009), the number of co-author papers and patents (Maggioni et al., 2007; Sonn and Storper, 2008; Matthiessen and Schwarz, 2010), the population of international migration (usually the business elite and senior) (Beaverstock et al., 2000), etc., and then builds a world city network system with the tools of graph theory and social network analysis method. However, nearly all the researchers acknowledged that the limitations of statistical data directly affected the accuracy and depth of the study.

In general, the two-mode dataset is characterized by connections between two separate sets of nodes (cities and firms, respectively) while actors in the one-mode dataset are directly interlinked. The essence of all corporate-based analyses of urban networks is a city-by-firm data matrix rather than a city-to-city adjacency matrix. Such researches have to face the transformation of the two-mode data into a one-mode data, and the transforming almost inevitably leads to a loss of information because certain network structures disappear after transforming.

Nowadays, the Internet has become an integral part of human economic and social life. The popularity of the Internet has made it possible to collect more detailed data, and greatly reduced the cost of data collection. With the arrival of the eras of cloud computing and big data, theoretical research about global urban network will break the past constraint of lacking connecting data, and embrace the new time where we have the ability to pay attention to each indivi
microscopically.

In the Internet era, people s online searching interest data of cities, to a certain extent, ca
comprehensively reflect the economy, image, openness, creativity and social vitality of cities., and has generated a growing impact on economic and social development. As a start for most users surfing online, web searching engine can find the desired content on the internet through a keyword. Therefore, the total number and frequency of keywords in the search engine can be a direct indicator of the degree of the concerned word.

In this paper, we take the one-mode dataset method and utilize the users concern data based on the world s largest search engine, Google, to characterize and reflect the mutual att
We expect to build a directed-weighted network of global cities based on the online inter-concern

data, and compare the final results with GaWC in order to reveal new features of urban network in the Information Age, also provide guidance for promoting the development of urban network.

2. Design of research methods

2.1 Data acquisition

To make it easier to compare our outcomes of with the predecessors' research achievements we take the 182 cities of three levels (Alpha, Beta, Gamma) according to GaWC 2012 as study subjects. With the key words index tool provided by Google Trends (Figure 1), we get the Google users search interest data of year 2012 among all 182 cities. Google Trends analyses a percentage of Google web searches to determine how many searches have been done for the terms users entered compared to the total number of Google searches done during that time. Trends results are normalized by the total searches from each region so that data from two regions with significant differences in search volumes can be compared equally. By writing scripts for batch query and file analysis, we build up an intercity concern matrix formed by 32,942 Trends data (see Table 1).

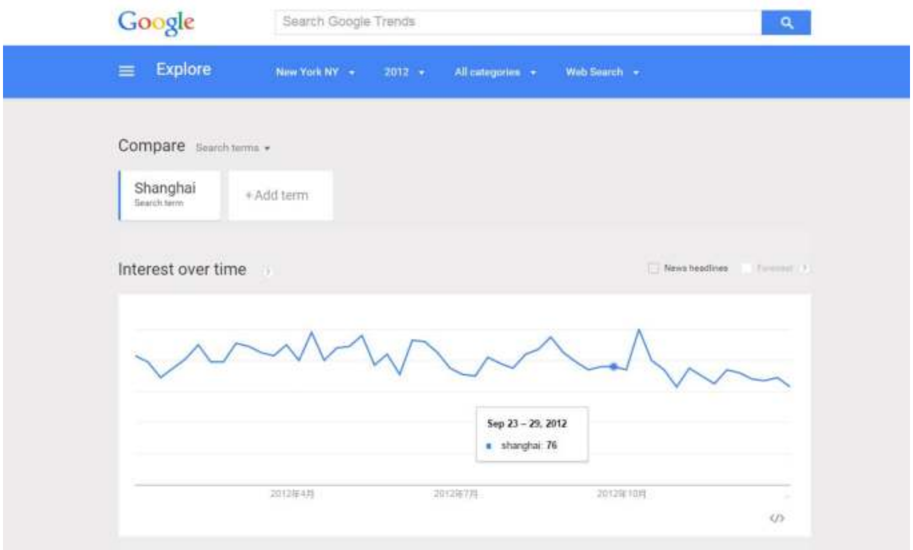

Figure 1. The screenshot of Google Trends page

Table 1. Pedagogical Matrix of Google Intra-searching Interest for Major World Cities

	City A	City B	City C	City D	Outdegree
City A		W_{AB}	W_{AC}	W_{AD}	OD_A
City B	W_{BA}		W_{BC}	W_{BD}	OD_B
City C	W_{CA}	W_{CB}		W_{CD}	OD_C
City D	W_{DA}	W_{DB}	W_{DC}		OD_D
Indegree	ID_A	ID_B	ID_C	ID_D	

To demonstrate the matrix, we take four cities A, B, C and D as an example (see Figure 2). There are two pieces of information link between each two cities. is the interest index of searching terms "city B" from users in city A. reflects to what extent people in city A concern about city B. Analogously, is the interest index of searching terms "city A" from users in city B. shows city A s popularity in city B. Thus, the matrix actually corresponds to the giant network of mutual searching links among world cities.

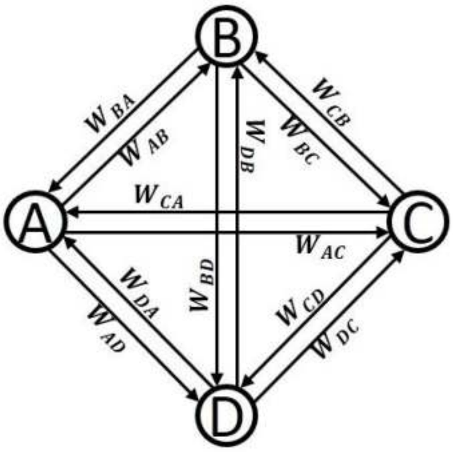

Figure 2. The Graphical Representation of Intra-searching Network

2.2 Data processing

Global urban network system can be studied from three aspects--**city nodes, information links, and urban networks.** This study starts from the data capture of the information links. Evaluation of the city nodes and the urban networks is conducted after processing the links.

2.2.1 Link

(1) Link strength

Each two cities has two reversed links between them (city A to city B and city B to city A). Using Taylor (2001) s model for reference, we define the strength of the link betwee multiplication of the two directional links. It is measured as follows:

$$= \quad = \tag{1}$$

The link strength index indicates the mutual concern degree between two cities. It is a comprehensive reflection of the strength of connectedness between the two cities. However, there is a problem with this algorithm: when the value of or is 0, the link strength index therefore equals 0, which is not consistent with the reality that there exists a one-way link between the two cities. To solve the problem, we learn from the method of calculating geometric mean, and improved the equation as follows:

$$= \quad = \sqrt{(\quad)^2 + (\quad)^2} \tag{2}$$

(2) Link symmetry

At present, the majority of the existing research literature of the global urban network focus on the non-directional weighted network. The importance of node and the directionality of flow are less clear and need to be taken into account. In this regard, Limtanakool et al. (2007) sets up a three S-dimensions framework for identifying and classifying the pattern of the urban systems from an interaction perspective. Limtanakool brings forward the concept of link symmetry index LSI to describe the degree of difference in both directions of the flow. For the link between cities A and B, LSI is measured as follows:

$$LSI_{AB} = LSI_{BA} = -\left[\frac{f\ \ln(f\ )\quad f\ \ln(f\ )}{\ln 2}\right] \tag{3}$$

$$_{AB} = \frac{W}{W \quad W} \tag{4}$$

$$_{BA} = \frac{W}{W \quad W} \tag{5}$$

Here, $_{AB}$ is the proportion of flow volume from nodes A to B in relation to the sum of volumes on links A to B and B to A. Similarly; $_{BA}$ is the proportion of flow volume from nodes B to A in relation to the sum of volumes on link A to B and B to A. When the value of LST_{AB} is 0, the link between nodes A and B is a one-way flow which means completely asymmetric. When the value of LST_{AB} is 1, the link between nodes A and B is an equivalent bi-directional flow, which means completely symmetric. Liu et al. (2013) thinks that Limtanakool s formula is inadequate in describing the vector nature of network links, because the index LST itself does not have directionality. LST is just limited to quantifying the symmetry level of the whole network. Furthermore, in the vicinity of 0.5 (where links A to B and B to A are close) LST does not have the sensitivity to reflect differences of the vectors. In order to realize the visual quantitative measurement of the symmetry of the two-way flow, Liu et al. make additions to the formula that when the volumes in both directions of the link have non-magnitude differences, the following formula can be used:

$$LST_{AB} = \frac{W \quad W}{W \quad W} \tag{6}$$

With Eq. (4), the formula can be evolved as follows

$$LST_{AB} = \quad _{AB} \quad 1 \tag{7}$$

When $LST_{AB} < 0$, $_{AB} <$ $_{BA}$, which means that node A s inflows are greater than outflows while node B s inflows are less than outflows. In the same way, when $LST_{AB} > 0$, $_{AB} >$ $_{BA}$, which means that node A s inflows are less than outflows while node B s inflows are greater than outflows. LST can distinguish the directionality of the flow by the positive and negative signs of the value, but the result is contrary to conventional understanding. Therefore, we further improve the formula as:

$$RLST_{AB} = 1 \quad _{AB} \tag{8}$$

When $RLST_{AB} = \quad 1$, flows completely goes towards node B from node A. When $RLST_{AB} = 0$, flows are equivalent on both directions between nodes A and B.

According to Eq. (2), (6) and (7), function diagrams are drawn to illustrate the relations of LSI_{AB}, LST_{AB}, $RLST_{AB}$ along with $_{AB}$ so that their differences in measuring city links are seen directly.

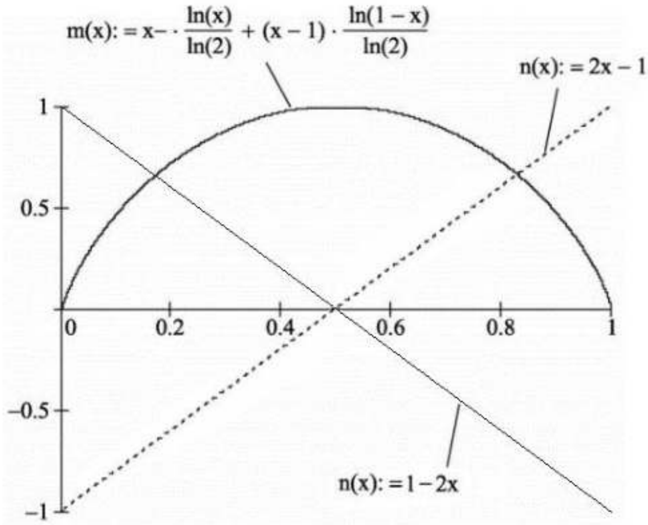

Figure 3. The Comparison of LSI, LST and revised LST index

2.2.2 Node

(1) Node strength

Node strength describes the sum of all link strength connected with a node. On the basis of Eq. (1), the node strength index of city A is measured as follows:

$$= \quad + \quad + \quad +$$

$$= \quad + \quad + \quad + \tag{9}$$

reflects city A s total amount of information flow in the global urban network.

(2) Node degree--indegree & outdegree

The node indegree index describes the total amount of information flowing into a node to a certain extent reflects the centrality of a city node. The node outdegree index describes the total amount of information flowing out of a node. to a certain extent reflects the radioactivity of a city node. For city A, its node indegree index and outdegree index are measured as follows:

$$= \quad + \quad + \quad + \tag{10}$$

$$= \quad + \quad + \quad + \tag{11}$$

(3) Node symmetry

We carry on the concept of node symmetry index NSI proposed by Limtanakool et al. (2007) to describe the degree of differences between node indegree and outdegree in the directional network. For node A, its node symmetry is measured as follows:

$$NSI_A = \frac{O}{O} \tag{12}$$

When $NSI_A = 1$, flows are completely going out of node A. When $NSI_A = 0$, inflows and outflows of node A is equivalent. When $NSI_A = 1$, flows are completely going into node A.

(4) Node Clustering Analysis

To compare with GaWC s study results, and identify differences and similarities of the world city hierarchies in various networks, this paper divides the 182 global cities into 3 classes and 10 sub-classes by K-means clustering taking the node strength index as variations. K-means clustering method is applicable to the case given a known number of categories. Its calculation principles are not discussed in this paper.

3. Empirical test

In this section, the proposed theoretical framework is illustrated by empirical data. The pattern of spatial interaction among 182 cities as constituted by Google Trends is characterized for a new world city network. Carrying out the test of these 182 cities enables us to assess whether our approach is feasible and effective.

3.1 Calculation of the link and node indices

Following the research methods above, we compute the values of all node and link indices of the network consisted of 182 world cities and their online inter-concern. Due to space limitations, detailed procedures are not included in this paper.

3.2 Comparison of Google Trends network and GaWC network

Does the new network based upon the web concern follow the old urban system? We make clustering analysis of the 182 cities, tabulate the outcomes along with GaWC s results (Table 2), and mark cities by classes on the world map (Figure 4).

Table 2. Values of Node Strength, Degree, Symmetry, Entropy and Class

Class		GaWC	Google Trends
	++	London; New York	Paris
	+	Hong Kong; Paris; Singapore; Shanghai; Tokyo; Beijing; Sydney; Dubai	Los Angeles; Chicago

		Chicago; Mumbai; Milan; Moscow; Sao Paulo; Frankfurt; Toronto; Los Angeles; Madrid; Mexico; Amsterdam; Kuala Lumpur; Brussels	Singapore; Toronto; Amsterdam; San Francisco; Frankfurt; Berlin; San Diego; Boston; Bristol; Birmingham(UK)
	-	Seoul; Johannesburg; Buenos Aires; Vienna; San Francisco; Istanbul; Jakarta; Zurich; Warsaw; Washington; Melbourne; New Delhi; Miami; Barcelona; Bangkok; Boston; Dublin; Taipei; Munich; Stockholm; Prague; Atlanta	Phoenix; Hong Kong; Charlotte; Atlanta; Vancouver; Dublin; Philadelphia; Edinburgh; Manchester; Washington; Dubai; Montreal; Rome; Zurich; Glasgow
	+	Bangalore; Lisbon; Copenhagen; Santiago; Guangzhou; Rome; Cairo; Dallas; Hamburg; Dusseldorf; Athens; Manila; Montreal; Philadelphia; Tel Aviv; Lima; Budapest; Berlin; Cape Town; Luxembourg; Houston; Kiev; Bucharest; Beirut	Seattle; Sydney; Vienna; Cleveland; Mumbai; Melbourne; Bangkok; Dallas; Stockholm; Istanbul; Copenhagen; Brussels; San Jose(US); Lyon
		Ho Chi Minh City; Bogota; Auckland; Montevideo; Caracas; Riyadh; Vancouver; Chennai; Manchester; Oslo; Brisbane; Helsinki; Karachi; Doha; Casablanca; Stuttgart; Rio De Janeiro; Geneva	Hamburg; Orlando; Leeds; Tampa; Portland; Calgary; New York; Oslo; Athens; Brisbane; Mexico; Bangalore; Manila; Sao Paulo; Seoul; Stuttgart; Cologne; Baltimore; Kuala Lumpur; Prague; Budapest; Moscow; Milan; Buenos Aires
	-	Guatemala City; Lyon; Monterrey; Panama City; San Jose(CR); Bratislava; Minneapolis; Tunis; Nairobi; Cleveland; Lagos; Abu Dhabi; Seattle; Hanoi; Sofia; Riga; Port Louis; Detroit; Calgary; Denver; Perth; Kolkata; San Diego; Amman; Antwerp; Manama; Birmingham(UK); Nicosia; Quito; Rotterdam; Belgrade; Almaty; Shenzhen; Kuwait; Hyderabad(IN); Edinburgh	Jakarta; Miami; Detroit; Chennai; Auckland; Pune; Hyderabad(IN); Dusseldorf; Rotterdam; Santiago; New Delhi; Perth; Marseille; Taipei; Milwaukee; Munich; Denver Johannesburg; Cape Town; Tel Aviv; Raleigh; Cairo; Geneva; Doha; Warsaw; Madrid; Adelaide; Lisbon; London; Columbus; Barcelona; Minneapolis; Belfast
	+	Zagreb; Lahore; St Petersburg; Jeddah; Durban; Santo Domingo; Baltimore; Islamabad; Guayaquil; St Louis; San Salvador; Cologne; Phoenix; Adelaide; Bristol; Charlotte; George Town(CI); Osaka; Tampa	St Louis; Rio De Janeiro; Riyadh; Kuwait; Lima; Utrecht; Seville; Shanghai; Colombo; Ankara; Beirut; Nairobi; Sofia; Karachi; Kolkata; Zagreb; Porto; Cincinnati; Antwerp; Lahore; Bucharest; Nantes; Krakow; Belgrade; Luxembourg; Leipzig; Turin; Hanoi; Bratislava; Jeddah; Lagos
		Glasgow; San Juan; Marseille; Guadalajara; Leeds; Baku; Vilnius; Tallinn; Raleigh; Ankara; Belfast; San Jose(US); Colombo; Valencia(SP); Cincinnati; Milwaukee; Muscat; Ljubljana	Casablanca; San Juan; Caracas; Abu Dhabi; San Jose(CR); Guadalajara; Ahmedabad; Kiev; Shenzhen; Kansas City; Guangzhou; Tunis; Beijing; Accra; Santo Domingo; Ljubljana; Monterrey; Quito; St Petersburg; Montevideo; Durban; Islamabad; Osaka; Panama City; Gothenburg; Ho Chi Minh City; San Salvador; Vilnius; Malmo

	-	Nantes; Tianjin; Accra; Algiers; Gothenburg; Porto; Columbus; Utrecht; Orlando; Ahmedabad; Asuncion; Kansas City; Seville; Turin; Dar Es Salaam; Portland; Krakow; Managua; Pune; Leipzig; Malmo; La Paz	La Paz; Guayaquil; Houston; Baku; Tokyo; Managua; Nicosia; Tallinn; Guatemala City; Asuncion; Dar Es Salaam; Almaty; Algiers; Manama; George Town(CI); Helsinki Muscat; Valencia(SP); Bogota Amman; Riga; Tianjin; Port Louis

From Table 3, we can see that:

1) In Class , 15 cities remain the same and 13 cities rank up compared with GaWC. Among them, the cities moving drastically up are Bristol, Charlotte, Phoenix, Glasgow, Edinburgh, Birmingham (UK) and San Diego. In addition, 4 cities--Paris, Singapore, Toronto and Hong Kong keep taking higher-ranking place in both two networks. Still more impressive, Paris climbs to the top from the 4th, Los Angeles soars to second from 18th, and Chicago leaps to third from 11th.

2) In Class , 30 cities remain the same; 14 cities rank up; 27 cities rank down. Among them, the cities moving drastically up are Portland, Orlando, San Jose (US), Pune, Leeds and Tampa, and the cities going largely down are London, Madrid, Barcelona, Johannesburg and Warsaw. Strikingly, London falls to 95th from the summit, and Madrid drops to 92th from 19th.

3) In Class , 14 cities remain the same and 42 cities rank down. Among them, the cities going largely down are Tokyo, Beijing, Bogota, Shanghai, Houston, Helsinki, Guangzhou and Ho Chi Minh City, most of which are located in the Asian regions. More prominently, Tokyo falls to 164th from 17th, Beijing drops to 143th from 8th and Shanghai slips to 107th from the 6th, which are the sharpest drop in all declines.

Thus, there exit significant differences between the new networks based upon inter-city concern on the Internet and the old urban system based on the producer service firms in cities.

Figure 4. World cities in 3 classes and 10 subclasses according to web concern

3.3 Visualization

This study involves massive data with non-uniform structures, so the visual data analysis tools (ArcGIS) is employed to map out the directional weighted urban networks formed by 182 city nodes and their inter-connected links (Figure 5).

In Figure 6, the size of city node responds to the value of the node strength index , indicating the city s sum volumes of searching and being searched. The lines correspond to the in the network where the value of the link strength index is shown by the line thickness. Nodes are divided into two types distinguished by colour: the red nodes are info-flow-in cities ($NSI < 0$) and the blue nodes are info-flow-out cities ($NSI > 0$). Arrows indicate the direction of flow.

According to the node symmetry index NSI, the 182 cities can be divided into two types: concerned-city type and concerning-city type. Concerned-city type refers to cities whose indegree is bigger than outdegree, which means the volume of searching for other cities is smaller than that of being searched. On the other hand, concerning-city type refers to cities whose outdegree is bigger than indegree, which means the volume of searching for other cities is bigger than that of being searched. In this study, the volume of being searched and that of searching for other cities are nearly equal in the case of most cities (where the value of NSI hovers near 0). A minority of cities manifest obvious directionality: the indegree of Helsinki, Riga, Muscat, La Paz and George Town (CI) is much bigger than outdegree so these cities obviously belong to the concerned-city type, while the indegree of Manama, Tokyo and Houston is much smaller than outdegree so these cities apparently belong to the other type. Very few cities present an extreme directionality: Bogota, Amman, Tianjin and Port Louis have a net inflow while Valencia (SP) has a net outflow. All cities showing obvious or extreme directionality belong to those whose total number of information flows are small and rank in the lowest class.

We sort every link between all the cities according to the flow volume and draw the results as a scatter plot (Figure 6). As can be seen clearly from figure 7, the values of the link strength index vary continuously from the minimum to the maximum, and there are 4 obvious stages throughout the function curve: 1) the value of is 0 where no connection happens among cities; 2) the value of rises rapidly from 0 to about 50; 3) the growth of slows down and forms a long gradual slope, where the majority of links are within this range; 4) the increase of accelerates again and tends to climb straightly, where the most important links with large volume of inter-concern are located. They are backbones of the whole urban network.

In order to more clearly understand the network structure, we display the map separately by setting top 5%, 10% and 20% of the biggest flow volume as threshold (Figure 7, 8 and 9).

When the threshold is set to top 5%, 10 core links appear in the world map. Among them, Birmingham (UK), Bristol and Manchester together form a strong inter-concern circle in Britain. Besides, the links Edinburgh-Glasgow, Nantes-Paris and Dusseldorf-Gothenburg further establish the heart position of Europe in the world cities network. For North America, two links, Cleveland-Cincinnati in the middle area and San Jose (US)-Francisco in the west coast, make America the second active region in the core network. For Southeast Asia, there is just one link from Hong Kong to Singapore visible. For Inter-continental connection, only one link from Zagreb to Calgary is visible.

When the threshold is set to top 10%, 49 links appear in the world map (including 10 core links). Compared with the situation of top 5% threshold described above, there are 6 more links turning up within the British Isles (Leeds-Bristol, Leeds-Birmingham, Leeds-Manchester, Glasgow-Belfast, Edinburgh-Belfast and Bristol-Dublin), 3 more links coming forth among other European cities (Paris-Berlin, Rotterdam-Amsterdam and Birmingham (UK)-Prague), 6 more links springing up in North America (San Jose (US)-San Diego (US), Washington-Baltimore, Columbus-Cleveland, Raleigh-Charlotte, Charlotte-Cleveland and Charlotte-Los Angeles, San Jose (CR)-San Jose (US)), 3 more links showing up in India (Pune-Mumbai, Ahmedabad-Pune and Hyderabad-Bangalore) and 2 more links showing up among other Southeast Asian (Singapore-Jakarta and Mumbai-Singapore)

When the threshold is set to top 20%, more than 800 links emerge in the world map. The embryonic framework has formed and is very close to the fully formed global urban network (Figure 5).

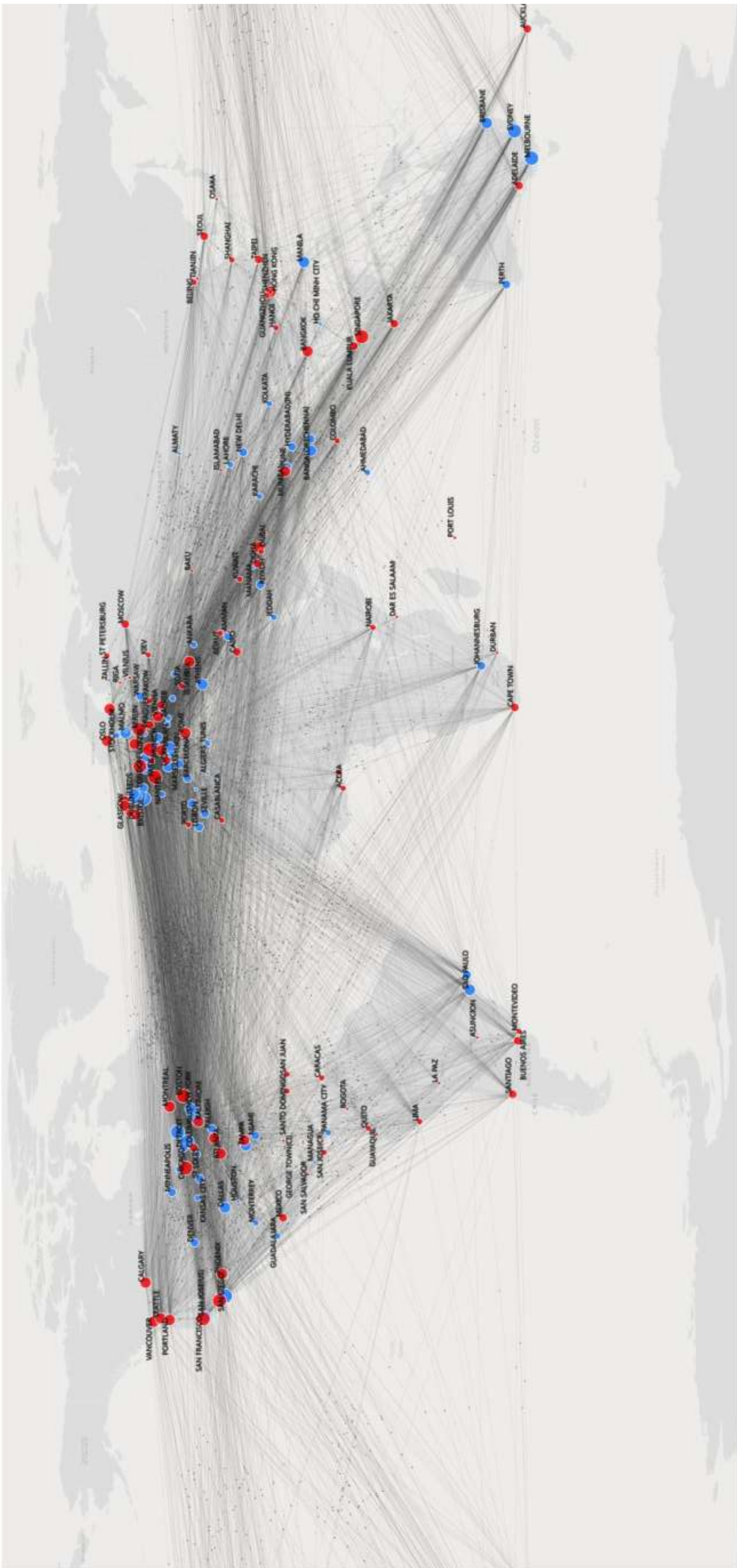

Figure 5. Map of the Global Urban Interconnectedness based on Google Trends

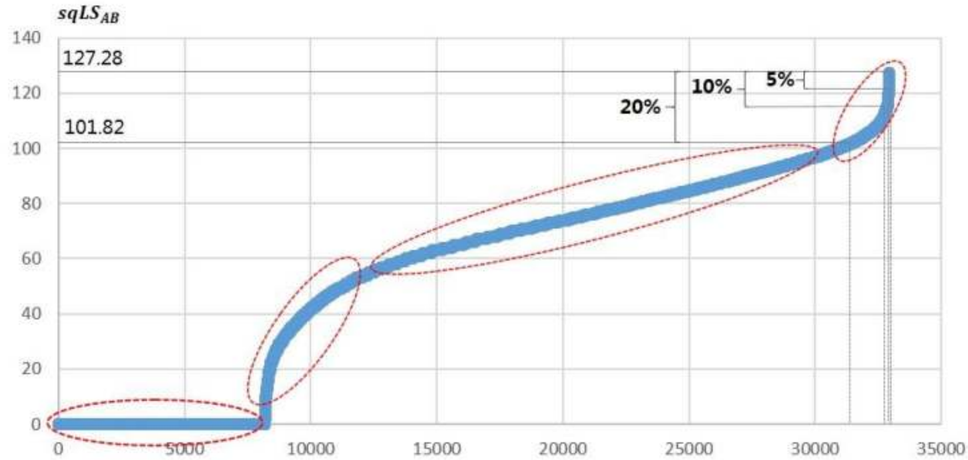

Figure 6. The scatter plot of

Figure 7. Intercity concern networks at threshold of top 5%

Figure 8. Intercity concern networks at threshold of top 10%

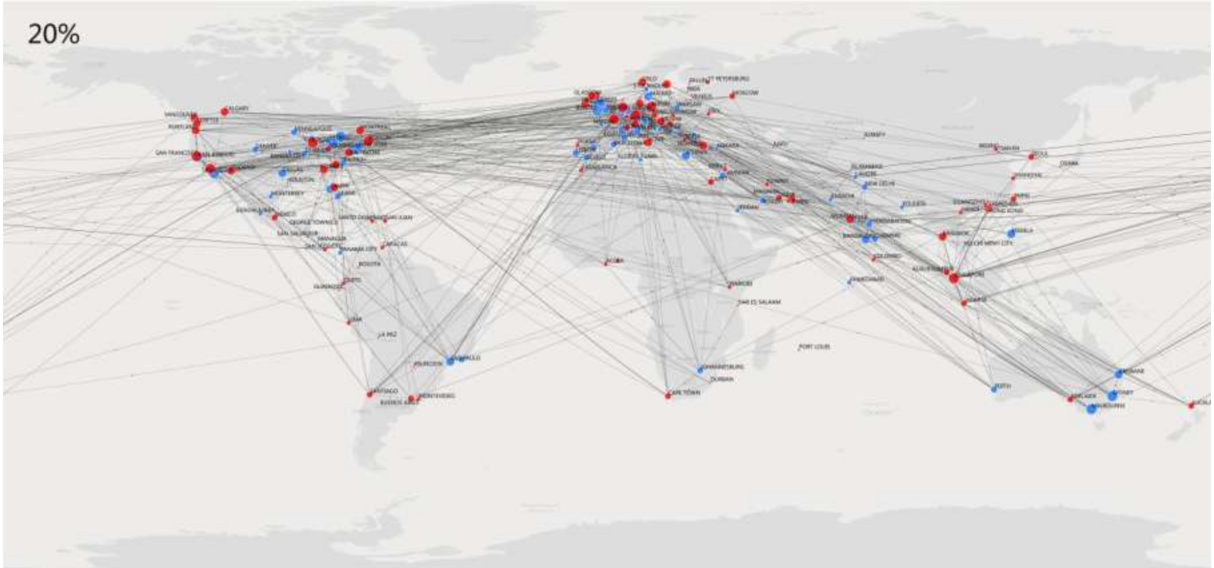

Figure 9. Intercity concern networks at threshold of top 20%

4. Conclusions and the way forward

Studies on global urban network have experienced three major phases: stage of world cities with attributes but without relations, stage of world cities in network society, overly new stage of world city network in global environment. In the third phase, analysis based upon one-mode dataset and two-mode dataset are the two main types. The well-known interlocking network model designed by Taylor (2001) for producer service firms in cities is of the two-mode dataset type. It has been widely adopted in academia, but the inherent problem of information loss caused by the projection from two-mode to one-mode makes it defective. Unlike two-mode type, one-mode dataset can directly create the city-to-city-matrix without transformation, while the persuasive and massive city-related data is needed.

In general, despite different underlying assumptions and operations, all approaches look at ways to build up the city-by- something datasets to estimate how individual cities are connected t the something networks. The rise of cloud computing and big data technologies provides us new possibilities for the something . In this context, we cast our eyes over the vast data of Internet users.

In this paper, we have (1) proposed a new framework with improved methods for analyzing the directed weighted network of city based upon one-mode dataset; (2) collected the Google users search interest data of year 2012 among 182 cities of three levels (Alpha, Beta, Gamma) according to GaWC; (3) simulated the information flow by calculating the link and node indices; (4) visualized the city network on geographical entities; (5) observed the hierarchies and patterns in the global city network; (6) made comparisons with GaWC group s outcomes.

Main results are as follows:

For **node**, 1) Europe has the largest number of higher-ranking cities in the new global urban network based on Google Trends, where the node strength is also greater, the next is North America, and the least is South America as well as Africa; 2) Southeast Asian cities except Singapore and Hong Kong all perform poorly under the new criteria; 3) Two types of cities, info-flow-in and info-flow-out, nearly distribute evenly across the main continents.

For **link**, 1) Europe still takes the heart position of the world, then North America and Asia. 2) Connection between Europe and North America is of the most density and frequency among other intercontinental links.

Overall, significant differences exist between the new networks based upon inter-city concern on the Internet and the traditional urban system based on the producer service firms. Interconnectedness of cities to a large extent is influenced by the real geographic location and distance. Cities in the same region are more likely to show concerns to their surrounding areas (the top 5% links in Figure 7 prove this). The district-crossing connection comes at the next level.

The analysing method of weighted network based on people s online attention da
flaws and advantages. The deficiencies and places might be improved of this article are as follows:

1) Although people s online attention data can comprehensively reflect the
openness, creativity and social vitality of a city, etc., the data s validity
recognized. The relationship between cities socioeconomic development and p
searching interest data still need to be verified through other inter-relational data between representative cities.

2) Trends data is relative, not absolute, which makes our map may not reflect the real scale of searching flows in the global context. Google s web page says that sets of search data are divided by a common variable, like total searches, to cancel out the variable's effect on the data. Although Google assures the comparability of Trends data, it still does not clearly explain how the raw data is normalized so that the restoration is unable to be carried.

3) English language bias. As a keyword for searching in this paper, the English names of the city may cause that the searching interest data of non-English-speaking areas cannot be truly reflected, while English-speaking world cities are highlighted in the network.

4) Although Google is the world's largest search engine, its service is limited in some countries and
regions, moreover, maybe it s not the mainstream search tool widely used by loca
cause that the data of Google Trends cannot fully reflect the real online attention.

5) From the perspective of graph theory, the node-related links, except the inflows and the outflows, also include the link whose source and destination nodes are the same. This kind of links, more universal and to some extent associated with city size, can indicate the contact between the various urban elements within the city. This paper focuses on the research of inter-concern network between
cities, ignoring the local people s search within the city itself, which may distort the netw
structures, especially the size of the nodes.

6) The specific reasons for the flow s asymmetry between cities need to be furth

7) This paper selects totally 182 cities distributed at , , three levels of the GAWC group’
results as study object, which has certain limitations. Besides, the distribution of these cities also cause some bias. For example, more European cities were selected compared to other regions and the online attention between these cities is far more than two distant cities, which makes these cities behave better than the actual situation.

The advantage lies in the analytical comprehensiveness of asymmetric links between cities and the technology of data collection and data processing, which saves us a lot of time.

1) As for analysis method, directional weighted network is an important supplement to the urban network research, which allows us to simultaneously analyse the directions and quantities of the
complex city network through studying flows asymmetry. This method helps us dee
the asymmetric of relationship between global cities and determine the true status of the various cities in the network more accurately. In the past, the research of global urban network can hardly go deep into the analysis level of directional weighted network because the data is difficult to obtain and handle. This study has provided technical support and empirical tests to remedy this defect.

2) As for analysing material, it s innovative to employ the statistical data ba this research, which has enlarged the studying range of original cyberspace in urban geography. The object of this study is measurable online data of bi-directional links between cities. Compared to the data of air flow, telecommunication flow, immigration flow, capital flow, etc., the online data are less expensive to obtain, easy to operate and convenient to use for urban and regional researchers. In this article only the Google Trends keyword search interest has been used, an offers other user's search-related indicators, which makes it potential for further research. Besides, with the improving of Internet data s abundance (such as commerce, education, t data), this approach may continue to enhance its own adaptability, which is also an important aspect in the follow-up studying.

Acknowledgement

We sincerely thank Li Ling, the technical director, for the constructive advice and the valuable support he has provided throughout the research process of this paper. The usual disclaimers apply.

References

- Beaverstock, J.V., Smith, R.G., Taylor, P.J., Walkera, D.R.F. and Lorimer, H., 2000. Globalization and world cities: Some measurement methodologies. Applied Geography, 20(1), pp.43-63.
- Castells, M., 1996. The Rise of the Network Society: The Information Age, Economy, Society and Culture. Volume I, Wiley-Blackwell, MA: Cambridge.
- Derudder, B. and Witlox, F., 2008. Mapping world city networks through airline flows: context, relevance, and problems, Journal of Transport Geography, 16(5), pp.305 312.
- Derudder, B. and Witlox, F., Faulconbridge, J. and Beaverstock, J., 2008. Airline data for global city network research: Reviewing and refining existing approaches. GeoJournal, 71(1), pp.5-181.
- Friedmann J., 1986. The world city hypothesis. Development and Change, 17, pp.69-83.
- Hall, P., 1966. The World Cities. London: Heinemann.
- GaWC. The World According to GaWC 2010. [online] Available at: <http://www.lboro.ac.uk/gawc/world2010t.html>
  [Accessed 5 January 2015].
- GaWC. Mapping Connectedness of Global Cities: , and tiers. [online] Availab ww.lboro.ac.uk/gawc/visual/globalcities2010.html>
  [Accessed 5 January 2015].
- Malecki, E.J., 2002. The economic geography of the internet s infrastructure Geography, 78(4), pp.399-424.
- Maggioni, M.A., Nosvelli, M. and Uberti T.E., 2007. Space versus networks in the geography of innovation: A European analysis. Papers in Regional Science, 86(3), pp.471-493.
- Matthiessen, C.W. and Schwarz, A.W., 2010. World cities of scientific knowledge: Systems, networks and potential dynamics. An Analysis Based on Bibliometric Indicators. Urban Studies, 47(9), pp.1879-1897.
- Limtanakool, N., Dijst, M. and Schwanen, T., 2007. A Theoretical Framework and Methodology for Characterising National Urban Systems on the Basis of Flows of People: Empirical Evidence for France and Germany. Urban Studies, 44(11), pp.2123 2145.
- Sassen, S., 1991. The Global City: New York, London, Tokyo. Princeton: Princeton University Press.
- Smith, D.A. and Timberlake, M., 1995. Conceptualising and mapping the structure of the world system s city system. Urban Studies, 32(2), pp.287-302.
- Smith, D.A. and Timberlake, M.F., 2001. World city networks and hierarchies, 1977-1997 An empirical analysis of global air travel links. American Behavioral Scientist, 44(10), pp.1656-1678.

- Sonn, J.W. and Storper, M., 2008. The increasing importance of geographical proximity in knowledge production: An analysis of US patent citations, 1975-1997. Environment and Planning A, 40(5), pp.1020-1039.
- Taylor, P. 2001. Specification of the world city network. Geographical Analysis, 33, pp.181 194.
- Taylor, P.J., 2004. World City Network: A Global Urban Analysis, Routledge, London.
- Tranos, E. and Gillespie, A., 2009. The spatial distribution of internet backbone networks in Europe. European Urban and Regional Studies, 16(4), pp.423-437.
- Xiong, L.F., Zhen, F., Wang, B. and Xi, G.L., 2013. The Research of the Yangtze River Delta Core Area s City Network Characteristics Based on Baidu Index. Economic Geography 33, No. 7, pp.67-73.