Open Geospatial Data,
Software and Standards

ORIGINAL ARTICLE   **Open Access**

CrossMark

Virtual network analysis of Wuhan 1+8 City Circle based on Sina microblog user relations

Shen Ying[1,2,3*], Nuocheng Zhang[1], Kai Wang[1], Zhigang Zhao[3], Zhou Shen[1] and Deguo Su[4]

Abstract

Background: In this study, we clustered a large number of microblog users into city nodes based on the interpersonal relations and geographic location attributes of these users and clustered the user relations within each city into the relations among cities in the Wuhan "1+8" City Circle (WCC) to transform and map the interpersonal network in the virtual network environment to an intercity network.

Methods: Based on this mapping, we analysed the characteristics proposed two evaluation indicators—degree of city cohesion and degree of city outreach—to quantify the relations within a city and the relations among different cities. The WCC urban network was analysed from a social network perspective, and the differences and similarities in a city's status between the virtual network and the real world were explored.

Results: As the "1" in the "1 + 8" WCC, Wuhan's decisive dominance in WCC was discussed from multiple perspectives to show the centre point of the circle and the monopoly status of Wuhan in WCC. This finding is consistent with the status of Wuhan in the actual geographic space as the political, economic and cultural centre of Hubei Province. Our study indicated tiny communication among the other 8 cities along the periphery of WCC in the virtual microblogging network space, with the exception of Wuhan; they closely resembled eight independent individuals.

Conclusions: The characteristics of the WCC intercity information exchange indicated that the concept of WCC did not stand out in the virtual microblogging network space and that these cities frequently exchanged information with cities outside WCC. And in WCC the degree of cohesion and outreach in a city node were significantly positively correlated with the economic development level of the city.

Keywords: Big data, Microblog user relation, Sina Microblog, Urban network space, Wuhan 1+8 city circle

Background

Cities interact and form relations under the influence of resources, industry distribution, information, population and other elements, which form urban network systems that are affected by various elements. The development and influence of cities have exceeded the scope and limits of these cities themselves, and cities and their key elements are closely connected by information networks, which form diverse global or regional urban networks (Zhen et.al, [1]). Urban networks have always been the focus of urban geography and economic geography studies, such as studies on networks in human settlements (including cities, towns and villages) that are formed by transport or logistics transport, such as rail and air transport (Wu et.al, [2]; Xue, [3]). Scholars are increasingly considering information technology as the driving force or factor of urban networks and focus their studies of urban and regional spatial relations on urban networks under the influence of informatization (Zhen et.al, [4]). For example, changes in a city's status in an urban network was analysed at different scales using Internet domain names and Internet Protocol (IP) addresses as indicators (Malecki, [5]; Rutherford et.al, [6]). Malecki

* Correspondence: shy@whu.edu.cn
[1]School of Resource and Environmental Sciences, Wuhan University, 129 Luoyu road, Wuhan 430079, China
[2]Ministry of Education, Key Laboratory of Geographic Information System (Wuhan University), Wuhan, China
Full list of author information is available at the end of the article

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 2 of 14

[5] and Townsend [7] analysed the network structures of the major cities in the world from the perspectives of Internet traffic and network bandwidth, respectively.

With the continuous development of Internet technology, the methods that people utilize to communicate are undergoing tremendous changes. The increasing rapid development of social network systems (SNSs) has significantly changed human social structure and activity patterns. Social networks that are based on certain social relations or common interests provide communication service for Internet users who gather online via these networks and extend human activities from the physical space to the virtual cyberspace (Kwan, [8]). Thus, social networks have integrated location-based services, which not only is more convenient for people to exchange information via social networks but also provides more methods for geography studies, such as the spatial structure of intercity information flow in virtual social networks, the relation system of urban spaces, spatial interaction mechanisms and urban spatial/temporal relationships. Social networks are critical for reconstructing and enriching theories of urban studies (Li et.al, [9]; Qin et.al, [10]) and can help social scientists reveal the mapping between the virtual cyberspace and the actual offline space, which compose the new organizational structure and activity patterns of human society (Yu and Shaw, [11]).

As a type of real-time broadcast social network platform that has rapidly developed and attracted an extensive range of users, microblogs broadcast information by "group" announcement. Thus, a microblog is a more appropriate tool for simulating urban social networks under the influence of informatization (Zhen et.al, [4]; Chen et.al, [12]). Due to the unidirectional relations among microblog users, users can build a network with coexisting strong relations and weak relations to satisfy their multi-level social needs (Ding et. al, [13]). The main directions for microblog data analysis are as follows: topic-event analysis (Ritter et.al, [14]; Yang et.al, [15]), sentiment analysis (Bollen et.al, [16]; Thewall et.al, [17]), information retrieval and recommendation (Spina et.al, [18]; Yao et.al, [19]), user network relation analysis (Recuero et.al, [20]; Yin et.al, [21]), information dissemination (Naaman et.al, [22]; Xue et.al, [23]), impact analysis (Ding et.al, [24]; Steeg and Galstyan, [25]), and analysis of intra-urban human mobility (Li et.al, [26]; Wu et.al, [27]). Li et al. [9] constructed China's intercity geo-social network based on the relations among Sina Microblog (SM) users and revealed that China's intercity social network was substantially imbalanced with significant global heterogeneity and local heterogeneity. Zhen et al. [4] analysed China's urban network hierarchy from the perspective of the microblog social network space and identified the centre cities of China's urban network and their clustering distribution. By associating a microblog

user's network information relations with his geographic location, Wang et al. [28] discovered that network information space was geographically rooted and manifested as local information relations dominate in the network information space and network information relations showed certain attenuation over distance. Based on the following relations of SM users, Chen et al. [29] employed flow analysis to demonstrate that China's urban network information relations had an asymmetric and spatially polarized structure with "one superpower and many" and observed certain similarities between the urban networks in the information space and the geographical distribution of areas with a high density of economic networks and transportation networks in the physical space. Chang et al. [30] clustered the social relations among individuals using users' location information on the city scale to construct an intercity social relationship network and presented a geographic backbone network extraction method that combines gravity models and information entropy technologies.
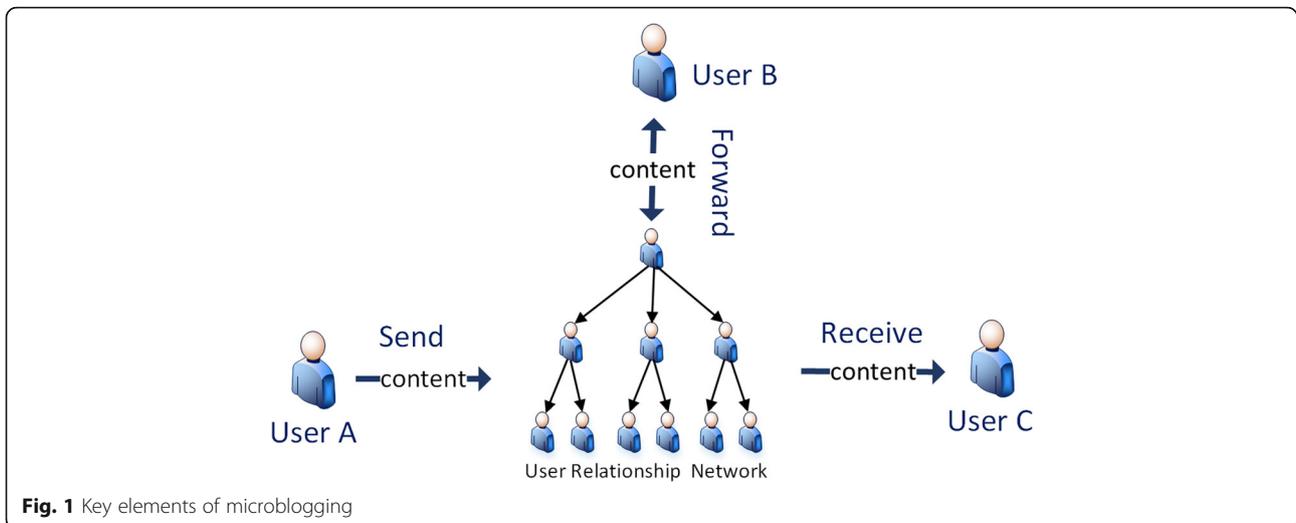
Based on the users' geographic location attributes and an analysis of relations among microblog users, we focused on the node of the network—the city—to analyse the relation strength within a city and among cities. Wuhan 1+8 City Circle (WCC) was employed as an example; its urban network characteristics were analysed from the perspective of a virtual social network to explore the similarities and differences in a city's position between the virtual network space and reality.

Methods

According to the data released by the Sina Microblog (Weibo) Data Centre, as of Dec. 2016 (Sina, [31]), SM had 297 million active microblog users, almost a quarter of China population. So we selected SM as the research platform and the WCC urban network as the target for our analysis.

SM user relations

Microblogging can form a complex network of relations via a large number of users and their relationships among them. The users can use the network to conduct real-time information dissemination and share and access information. Microblogging has three features: the real-time transmission of information, the real-time transmission of concise content and user interactions. The three key components of microblogging are as follows (Wang et.al, [32]): the users, microblogging content and user relations, and the functions of the user relationship network are reflected in three aspects: sending, forwarding, and receiving, which similar with Twitter or Facebook (Fig. 1). In a microblog, the relations among users are not mutual and asymmetric. The traditional mutual relationship is divided into two separate parts: the user can follow a user but does not always require

Ying et al. Open Geospatial Data, Software and Standards (2017) 2:2

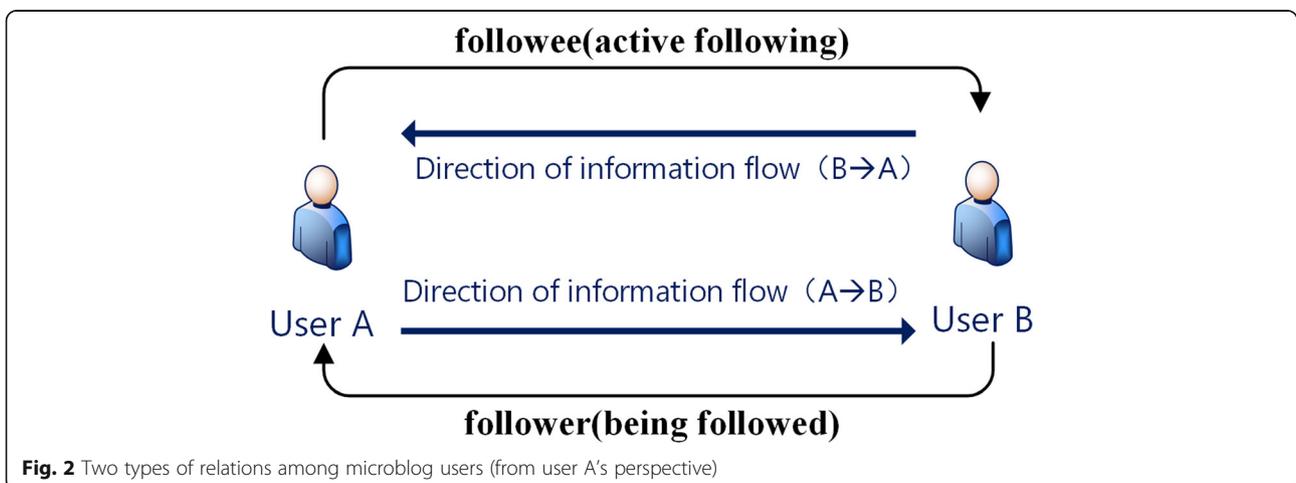Page 3 of 14
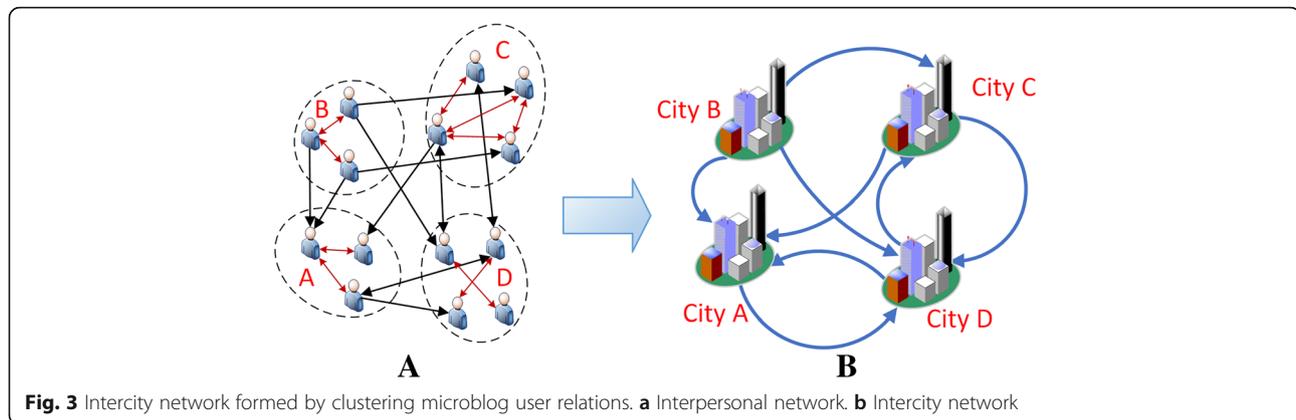

Fig. 1 Key elements of microblogging

the user to follow himself/herself (Golder and Yardi, [33]); indirect and closed ternary relations may exist (Boyd et. al, [34]). Microblog user relations in a virtual world comprise purely interpersonal networks; the relations in this network may consist of either "strong ties" (relations established between acquaintances) or "weak ties" (relations established between strangers) (Granovetter, [35]). The interpersonal relations in this virtual world enable information flow. The information flows among microblog users are transmitted by the complex networks that are formed by following users' relationships, which can be divided into two types—"followees" and "followers", that's the user follows others actively and be followed by others passively, as shown in Fig. 2.

According to user registrations and login location attributes, a large number of microblog users can be clustered as city nodes based on their city's geographical location. Therefore, the user relations within each city can be clustered as relations among cities, which can be utilized to map and transform interpersonal networks in a virtual network environment to intercity networks.

SM has hundreds of millions of users who are located in various provinces and cities nationwide; users' locations are stored as attributes in their user profiles. Therefore, a city can be treated as a node to cluster users of the same city. Microblog user networks are similar to other graph-based networks, and the users and relations among users can be abstracted as points and lines in a graph according to graph theory (Fig. 3a). City nodes contain a large number of microblog users, and the relations among cities are determined by the amount of following relations among these microblog users. In urban networks, cities and the relations among cities can be abstracted as points and lines in a graph using graph theory (Fig. 3b). These points and lines differ from but are also related to their counterparts in a microblog space. For example, users inside the same dotted frame in Fig. 3a originate in the same city, which

Fig. 2 Two types of relations among microblog users (from user A's perspective)

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 4 of 14

Fig. 3 Intercity network formed by clustering microblog user relations. **a** Interpersonal network. **b** Intercity network

is clustered as a node in Fig. 3b. The intercity relations, whose strength is determined by the relations among their microblog users, are also formed. We employ this information to construct and map an intercity network from interpersonal networks from the user connection level, and there is on space in this paper to discuss the contents of Weibo, the frequency and volume of information flow from detail level. In order to reflect the actual urban community network environment of the entire city circle synthetically from the holistic view of Weibo users, we do not consider user information (e.g. gender, age, education, and employment) or individual views.

Study area and data

WCC was officially approved by the State Council of China in December 2007. WCC refers to the group of cities within 100 km from the largest city in central China—Wuhan—including eight surrounding medium-sized cities: Huangshi, Ezhou, Huanggang, Xiaogan, Xianning, Xiantao, Tianmen and Qianjiang. WCC covers less than a third of Hubei Province. However, WCC contains half of the province's population and produces more than 70% of the provincial gross domestic product (GDP). WCC is one of the largest city groups in central China. Wuhan city is the centre of the circle, Huangshi is the second centre city and Xiantao is the centre city in its west wing (Fig. 4). There have already many researches about realistic analysis about land use and urban spatial expansion (He et.al, [36]), road network (Liu et.al, [37]) and field source-strength method for interaction scenario (He et.al, [38]).

SM user samples were selected from each WCC city and the user city is determined based on the location attribute on the user profile, whether they use SM in mobile environment or computer platform. The data include two parts: personal information about individual microblog users and the user relation data. Each user sample corresponds to multiple followers and followees who he/she follows. The conditions for user screening are as follows: 1) the user is located in WCC; 2) the user sample contains both "grassroots" and "celebrities" to reflect the real-world network environment, and "zombie fans" are excluded according to the definition: invalid accounts signed up by network companies for the purpose of increasing the number of fans for certain users, especially celebrity ones, and getting them more attention (Yuan et. al, [39]); and 3) the number of user samples for each selected city is proportional to the total number of microblog users of that city to reflect the actual urban community network environment of the entire city circle. According to these requirements, we employed the SM open platform application programing interface (API) and traditional web parsing technologies (Lian et.al, [40]; Zhou et.al, [41]) to develop a web crawler program to collect data. A total of 31,481 valid random user samples (including 16,680,418 followers and 10,102,556 followees) were collected (See Additional file 1). All data were collected between July 2, 2015 and August 17, 2015. Table 1 lists detailed information about the user samples that were collected in the experiment. We had to assume that random sample can represent the whole uses, because random characters and the proportion of the sample size to the whole population in each city is approximately equal.

Degree of city cohesion and city outreach

To quantify the characteristics of urban network information dissemination in urban nodes, two indicators—city "cohesion degree" and city "outreach degree"—were proposed to describe the information dissemination distribution in the urban network for each city. The following relationships among microblog users have directions, which can be divided into two types of relationships (followee and follower)

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 5 of 14

Fig. 4 Map of the Wuhan "1+8" City Circle

depending on the following direction. To reflect the direction of the following relationship (i.e., the opposite direction of the information flows) in the information transmission features of the city node, we separately calculated the degree of cohesion and outreach of the followers and followees.

The degree of follower cohesion ($N_i^{'}$) characterizes the proportion of sampled users' followers within the same city to all their followers, and the degree of followee cohesion ($N_i^{''}$) represents the proportion of sampled users' followees within the same city to all of their followees. The corresponding formula is expressed as

$$N_i^{'} = U_{ii}^{'} / \sum_j U_{ij}^{'} \tag{1}$$

$$N_{i''} = V_{ii}^{'} / \sum_j V_{ij}^{'} \tag{2}$$

where $U_{ii}^{'}$ is the normalized follower count of city i within city i, $V_{ii}^{'}$ is the normalized followee count of city i within city i, $\sum_j U_{ij}^{'}$ is the normalized total follower count of city i within all cities, $\sum_j V_{ij}^{'}$ is the normalized total followee count of city i within all cities, and j is the number of selected cities.

To be precise, the "follower count" of a city is the sum of the followers of all individual nodes in that city, but followers count is divided into two parts (inner-city and outer-city followers) in latter calculations. As in Equations 1 and 4, $U_{ii}^{'}$ means inner-city followers of city i while $\sum_{j,j \neq i} U_{ij}^{'}$ means outer-city followers of city i, and their sum $\sum_j U_{ij}^{'}$ means the "follower count". For "followee count", it is similar. These processes can eliminate the influence of inner-city relationships and make the quantitative description between the cities more precisely.

Table 1 Basic information about the user sample sets

City	User sample size	Follower counts	Followee counts
Wuhan	27603	15068908	8916992
Huangshi	1043	482506	290835
Ezhou	311	126373	120871
Huanggang	827	306626	257585
Xiaogan	784	338722	246409
Xianning	546	209708	184039
Xiantao	154	64483	30140
Tianmen	94	36318	30082
Qianjiang	119	46774	25603
Total	31481	16680418	10102556

Combining the degree of follower and followee cohesion, the degree of city cohesion (N_i) becomes a comprehensive representation of the closeness of the current city as it relates with itself. The degree of city cohesion reflects the characteristics of mutual following and is calculated as follows:

$$N_i = \left(U'_{ii} + V'_{ii} \right) / \left(\sum_j U'_{ij} + \sum_j V'_{ij} \right) \qquad (3)$$

Similarly, the degree of follower outreach (M'_i) and followee outreach (M''_i) of certain city can be calculated as:

$$M'_i = \sum_{j, j \neq i} U'_{ij} / \sum_j U'_{ij} \qquad (4)$$

$$M''_i = \sum_{j, j \neq i} V'_{ij} / \sum_j V'_{ij} \qquad (5)$$

Accordingly, the degree of city outreach (M_i) describes the distribution of the city's information flow

direction in the urban network. The degree of city outreach represents the proportion of outer city followers and followees among all followers and followees of the city's microblog users. The corresponding formula is expressed as

$$M_i = \left(\sum_{j, j \neq i} U'_{ij} + \sum_{j, j \neq i} V'_{ij} \right) / \left(\sum_j U'_{ij} + \sum_j V'_{ij} \right) \qquad (6)$$

where $\sum_{j, j \neq i} U'_{ij}$ is the normalized outer city follower count of city i and $\sum_{j, j \neq i} V'_{ij}$ is the normalized outer city followee count of city i.

According to Equations 3 and 6, the degree of cohesion and outreach of the followers or followees of the same city sum to 1.

The follower cohesion, followee cohesion and city cohesion were calculated for each city according to these formulas; the results are shown in Fig. 5a. The degree of city cohesion are sorted and arranged in descending order along the horizontal axis.
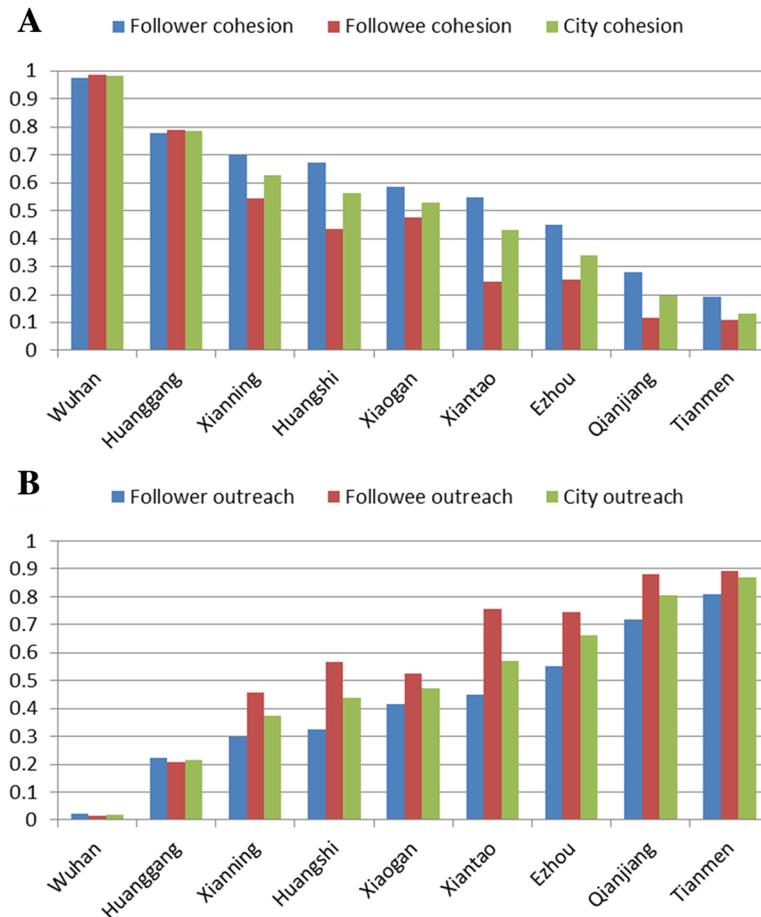
Fig. 5 Comparison of the Degree of Cohesion and Degree of Outreach in WCC Cities. **a** Comparison of the three types of degrees of cohesion among nodes in urban networks. **b** Comparison of the three types of degrees of outreach among nodes in urban networks

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 7 of 14

Figure 5a shows that the degree of city cohesion in Wuhan was significantly larger than the degree of city cohesion for the other eight cities, which indicates that microblog users in Wuhan were relatively more inclined to exchange information with inner-city users and that inner city microblog users are closely related with each other while regarding the WCC as a closed circle. This finding also reflects that Wuhan was very appealing and could attract inner-city users and that the degree of acceptance toward other cities was not significant in Wuhan. Thus, Wuhan is a more biased and closed city in the information space with a strong urban culture cohesion and a relatively lower level of openness and inclusiveness.
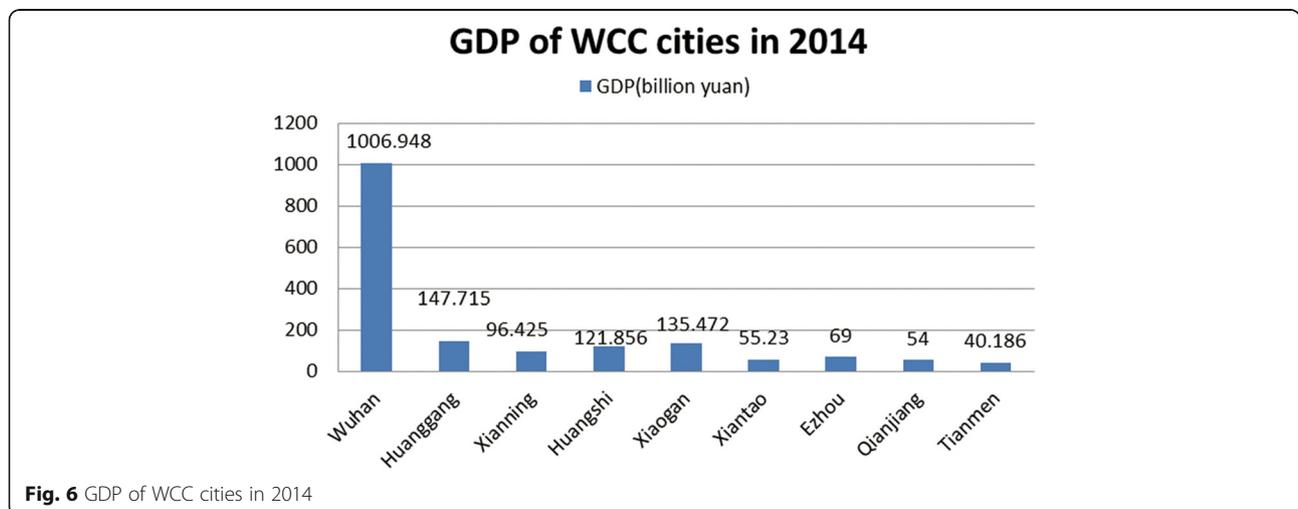
Qianjiang, Tianmen and other cities showed lower degrees of cohesion, and their inner-city microblog users were more likely to exchange information with outer-city microblog users. Generally, these cities had less influence and a smaller voice in the information network space, and inner-city users frequently lacked the ability to attract each other. In microblogging communities that are dominated by weak relations, the motivation to exchange information inside these cities was frequently insufficient, and users were more likely to choose other cities for information exchange.

As shown in Fig. 5, the followee cohesion degree in Wuhan and Huanggang was greater than the corresponding follower cohesion degree, which suggests that inner-city microblog users connected more to other inner-city microblog users and receive information updates. This finding also reflects that Wuhan and Huanggang were more appealing and had a stronger voice in WCC. For all other WCC cities, the degree of followee cohesion was lower than the degree of follower cohesion, which reflects that the microblog users in these cities were more willing to receive information from microblog users in other cities instead of receiving updates from microblog users in the same city.

By definition, the degrees of cohesion and outreach are complementary for each city; thus, the proportion of users that are connected with outer-city users decreases as the proportion of users that are connected with inner-city users increases, and vice versa. Fig. 5b shows the comparison of the degree of outreach among different node cities in the urban network. Fig. 5b indicates that Wuhan had the lowest degree of outreach, which suggests that information was primarily exchanged within the city but rarely with other cities. Tianmen, Qianjiang, Ezhou, Xiantao and other cities ranked high regarding the degree of outreach, which suggests that microblog users in these cities primarily exchanged information with users in other cities but rarely with users of the same city.

When comparing the ranking in Fig. 5a with the GDPs of these cities, we observed an interesting correspondence. The GDP data for all WCC cities were retrieved from the National Bureau of Statistics of China (as shown in Fig. 6). The simple comparison of Figs. 5a and 6 indicates that the GDP ranking was highly positively correlated with the degree of city cohesion and highly negatively correlated with the degree of city outreach. Potentially, a city's attractiveness and word power were significantly influenced by its economic factors: the attractiveness and word power and the influence increased as the economic conditions increased. Cities with a higher GDP were more likely to engage in information exchange inside the city and tended to be more conservative and closed in the virtual information space. In contrast, cities with relatively low GDPs were more inclined to engage in information exchange with cities with higher GDPs and tended to be more open and inclusive in the information space.

Fig. 6 GDP of WCC cities in 2014

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 8 of 14

For each city in an urban network, the population, transportation or information flow can be divided into two parts: the exchange that occurs inside the city and the exchange with other city nodes in the urban network. The concepts of degree of city cohesion and degree of outreach that are proposed in this paper reflect the information flow characteristics of a city in microblog space. These characteristics can reflect the information flow characteristics of the city and, from the microblog perspective, the intensity of the information exchange with other city nodes. Microblog users of cities with a higher degree of cohesion were more willing to exchange information with other microblog users in the same city, and information frequently flowed inside the city. Conversely, a higher degree of outreach suggests that the city had a close relation with other cities. The degree of city outreach reflects the closeness and frequency of intercity information exchange, which is a new perspective and method for reflecting the characteristics of intercity relations.

Results and discussions

City outreach strength analysis

To reflect the strength of the relations among various city nodes, the closeness of the relations among cities should be analysed. The strength of the relation between city i and other cities can be calculated using the following equation:

$$P_{ij} = \left(U'_{ij} + V'_{ij} \right) / \left(\sum_{j \neq i} U'_{ij} + \sum_{j \neq i} V'_{ij} \right) (i \neq j) \qquad (7)$$

where P_{ij} represents the outreach strength of city i towards city j, U'_{ij} represents the normalized number of followers of microblog users in city i who are located in city j, and V'_{ij} represents the normalized number of followees of microblog users in city i who are located in city j. $\sum_{j \neq i} U'_{ij}$ represents the normalized total number of outer-city followers of microblog users in city i, and $\sum_{j \neq i} V'_{ij}$ represents the normalized total number of outer-city followees of microblog users in city i.

Based on the follower-followee relationship among SM users, their geographic information and the number of users, the strength of the relations of various cities with other cities can be calculated using these formulas, namely, the strength of the relation between a city and other cities.

A Nightingale's Rose Diagram was plotted based on the strength of the intercity relation of the concerned city, as shown in Fig. 7.

In the series of Nightingale's Rose Diagrams, the relation strength increased as the radius and area increased. This type of rose diagram can accurately and intuitively show the relation strength of the city concerned with other cities. With the exception of Wuhan, the relation

with Wuhan contributes to 90% of the relation strength for the other eight cities. The relation fully reflected the decisive dominance of Wuhan City in the information space, namely, the "1" in the "1 + 8" WCC and the mega city in the city circle. The relation with the other eight cities (excluding Wuhan) accounted for less than 10%, which indicates a very weak relation. The communication among these 8 cities was minimal, and the vast majority of their communication interacts with Wuhan. Fig. 7a shows the relations between Wuhan and the other eight WCC cities, and the strength of the relation with Wuhan in descending order were as follows: Huanggang, Xiaogan, Huangshi, Xianning, Ezhou, Xiantao, Qianjiang, and Tianmen. This order is exactly the same order as the GDP ranking shown in Fig. 6 (ranking: Wuhan, Huanggang, Xiaogan, Huangshi, Xianning, Ezhou, Xiantao, Qianjiang, and Tianmen). Therefore, they were highly positively correlated with each other, which suggests that the strength of Wuhan's relations with cities inside WCC is substantially dependent on the economic factors of the city. The results indicate that a city with a strong economy will attract a significant amount of attention.

Intercity relation in WCC

The intercity relations are formed by the distribution of followers and the mutual following. As different cities considerably differed in terms of the number of followers and followees, we normalized the sum of the data that were collected to eliminate the impact of absolute count differences on the final results. The following formula can be applied to normalize the follower count data.

$$U'_{ij} = U_{ij} / \sum U_{ij} \qquad (8)$$

where U'_{ij} is the normalized number of followers of city i who are from city j, U_{ij} is the number of followers of city i who are from city j, and $\sum U_{ij}$ is the total follower count of city i. Similarly, the followee count data can be processed with the following formula:
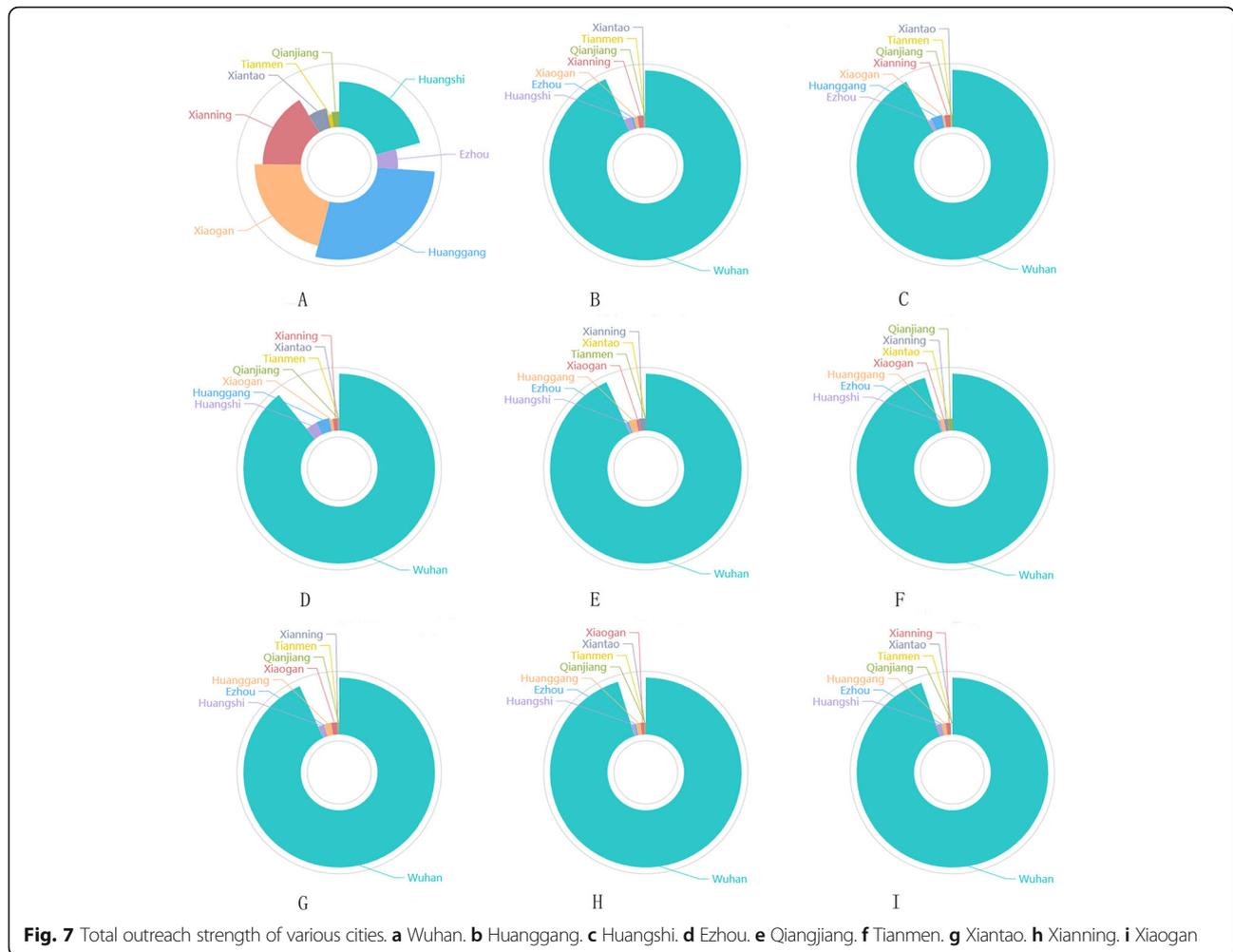
$$V'_{ij} = V_{ij} / \sum V_{ij} \qquad (9)$$

where V'_{ij} is the normalized number of followees of city i that are from city j. V_{ij} is the number of followees of city i that are from city j, and $\sum V_{ij}$ is the total followee count of city i.

In addition, the total intercity relation strength can be expressed with the following formula based on both follower and followee data:

$$\mathrm{W} = \left(U_{ij} + V_{ij} \right) / \left(\sum U_{ij} + \sum V_{ij} \right) \qquad (10)$$

According to equations (8) to (10), the strength of the three types of intercity relations were calculated and shown in Figs. 8a-c.

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 9 of 14

Fig. 7 Total outreach strength of various cities. **a** Wuhan. **b** Huanggang. **c** Huangshi. **d** Ezhou. **e** Qianjiang. **f** Tianmen. **g** Xiantao. **h** Xianning. **i** Xiaogan

As shown in Fig. 8, Wuhan as "1" in the "1 + 8" WCC paid more attention to itself and minimal attention to the remaining eight cities. In terms of the information exchange of the other eight cities, the strength of the relations towards Wuhan and towards themselves accounted for a considerable part of the total relation strength; their combination accounted for approximately 95%, whereas their connection with the other seven cities (excluding Wuhan and the city itself) did not exceed 5%. However, the situations of these eight cities (excluding Wuhan) also differed (As shown in Fig. 9). For example, Huanggang's strength of relation with itself accounted for nearly 80%, whereas Qianjiang and Tianmen's strength of relation with itself accounted for less than 20%. This phenomenon is likely related to local economic factors and cultural factors. With the exception of Wuhan, in WCC, a city with a higher GDP rank would have stronger relations with itself and relatively weaker relations with the other seven cities (excluding Wuhan and the city itself) based on the GDP rankings. These results also verified the earlier observation that the stronger the economic strength is, the greater the attention (including attention from itself) that it attracts.
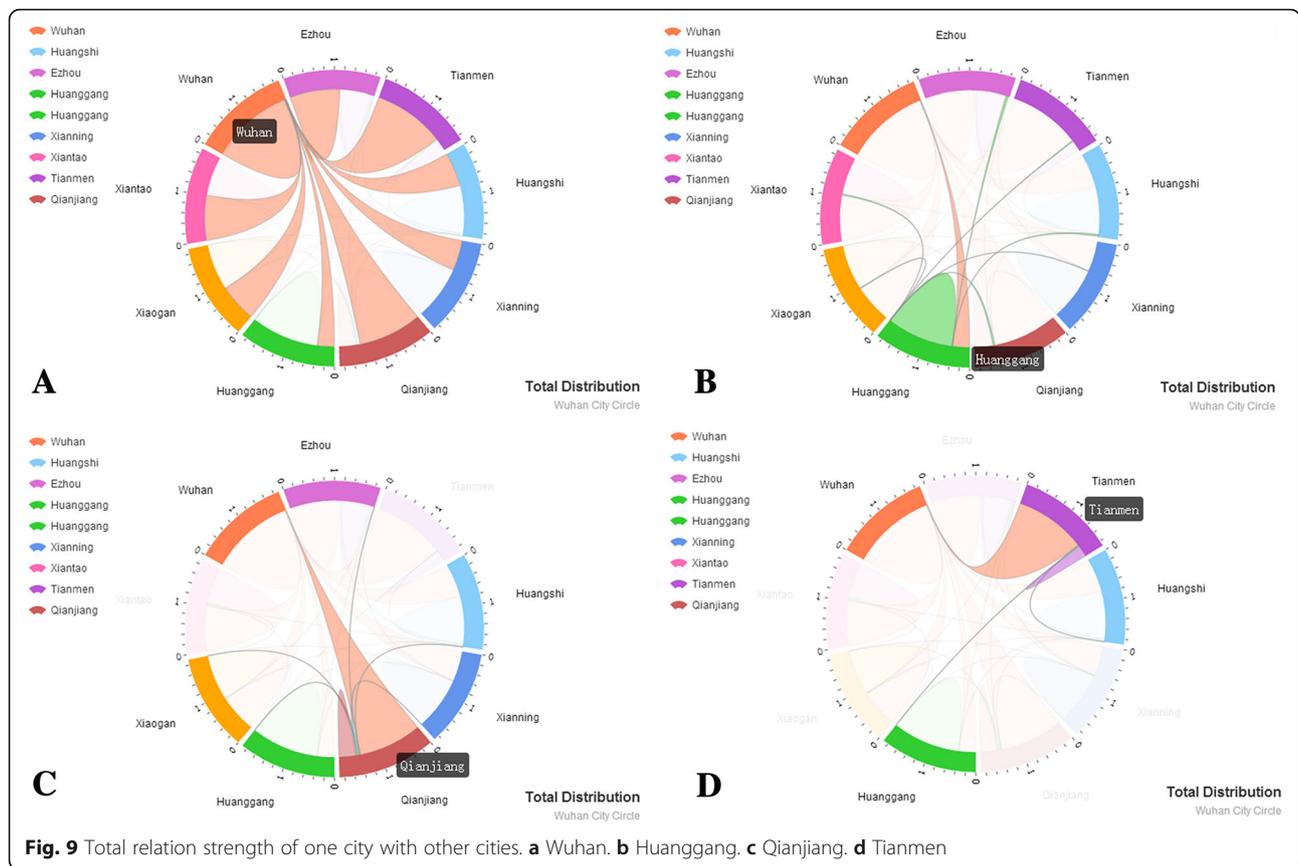
In the actual geographical space, WCC centres around Wuhan with Huangshi as the second centre city and Xiantao as the centre of its west wing. In the virtual cyberspace, if the eight cities, with the exception of Wuhan, are treated as a new circle in the analysis, the visualization of the follower-followee data would indicate that each city in this new city circle primarily exchanges information within itself with minimal communication with other cities (As shown in Fig. 10). This new circle does not seem to have a centre city; instead, it has eight independent and equal individuals. In the virtual cyberspace, Huangshi did not have the status as "the second centre city" in WCC, and Xiantao distanced itself as "the centre city in the west wing". This finding was inconsistent with the geospatial reality, which indicates a certain level of difference between the virtual cyberspace and the geospatial reality.

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 10 of 14
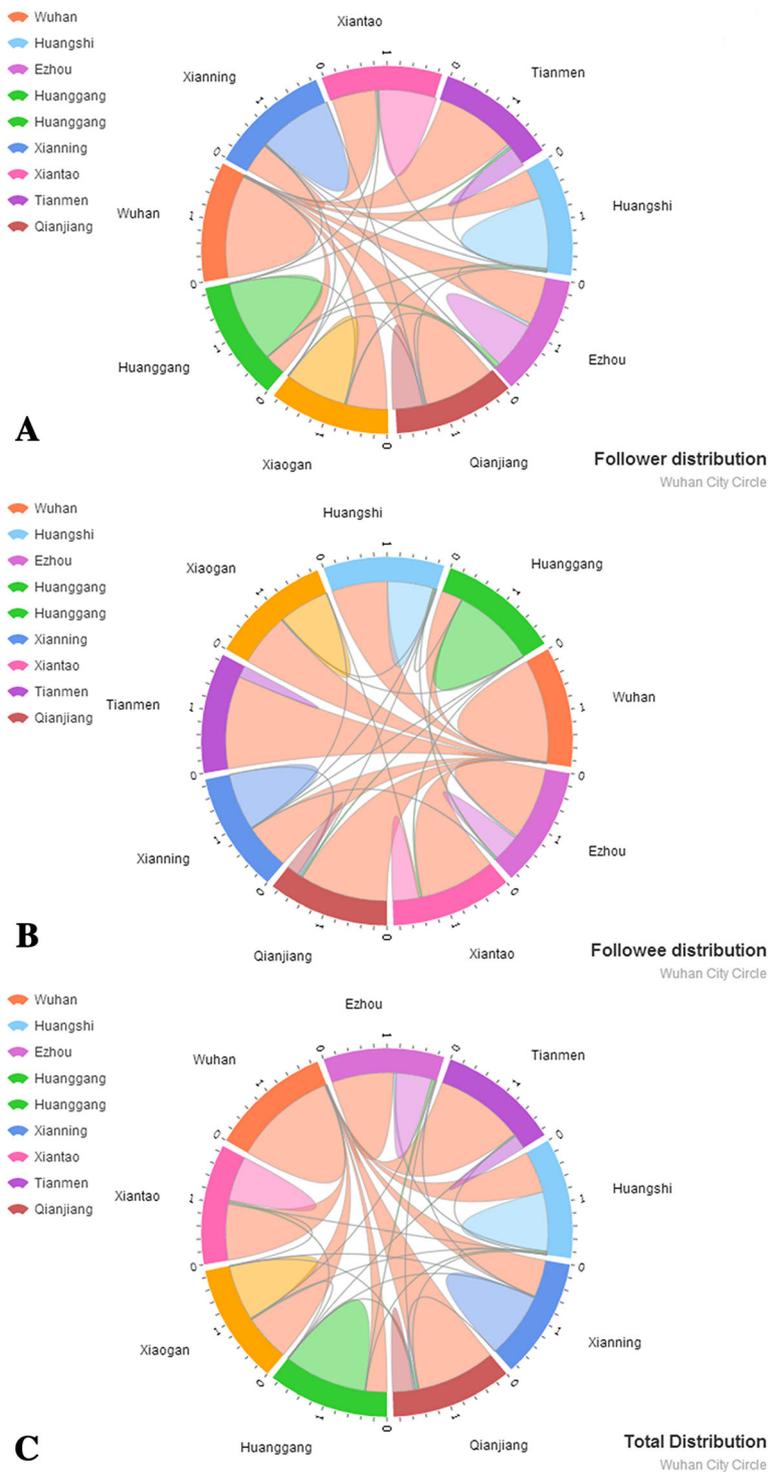
Fig. 8 Distribution of intercity relation strength of various WCC cities. **a** Follower distribution. **b** Followee distribution. **c** Distribution of total intercity relation strength

And we underline this works only within the study area WCC. This might not work if the geographic area is extended to the whole China. We also investigate connections of cities in the WCC and outside the WCC

particularly national or regional large cities. Since there are 34 provincial-level cities (e.g. Wuhan, Beijing, Guangzhou), 333 prefecture-level cities (e.g. Huangshi) and 2862 country-level cities, Wuhan 1+8 City Circle

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 11 of 14

Fig. 9 Total relation strength of one city with other cities. **a** Wuhan. **b** Huanggang. **c** Qianjiang. **d** Tianmen

including 9 cities is a very small part in China city network (9/3229). And the connections of cities in the WCC are also quite weak in China city network, which will significantly impair the connections inside the WCC. And the follwers and followees of WCC only occupy 1.5 and 3.3% respectively inside WCC. Whether the cities inside WCC are very likely to have strong connections with cities outside the WCC (like Beijing) is out of the scope of this paper.

Conclusion

Although geographic methods can be employed in the field of geography from the perspective of information technology, new issues and challenges arise. The following relationships of SM users determine the flow of information dissemination among users and were applied as the starting point and incorporated into the urban network analysis in this study. We analysed this type of intercity following relationship from a new perspective to reveal the structural characteristics of an urban network by clustering the following relationships among microblog users into the following relationships among cities. The degree of city cohesion and outreach are depictions of the characteristics of urban information flow from the perspective of

microblog space. Due to differences in geographical location, cultural attitudes, proportions of foreign population and other factors, users in different cities will consciously or unconsciously exhibit their characteristics and differences in information exchange. Therefore, the information exchange characteristics of a city can be obtained by analysing microblog indicators.

Study regions varying from local city (e.g. Wuhan) to city group (e.g. WCC), to whole China, will gain different views and insights. Only cities in WCC and their relations were compared in this study; the influences of other cities in China were neglected. The massive data of SM users can ensure the accuracy of the study and the reliability of the results analysis. From the individual level, we gain the network quantitative analysis and relationships at the city level in WCC. By analysing the urban network structure, we obtained the characteristics of the nodes and links of WCC in the virtual microblogging network space:

- In this paper, as the "1" in the "1 + 8" WCC, Wuhan's decisive dominance in WCC was discussed from multiple perspectives to show the centre point of the circle and the monopoly status of Wuhan in WCC. This finding is consistent with the status of Wuhan in

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 12 of 14
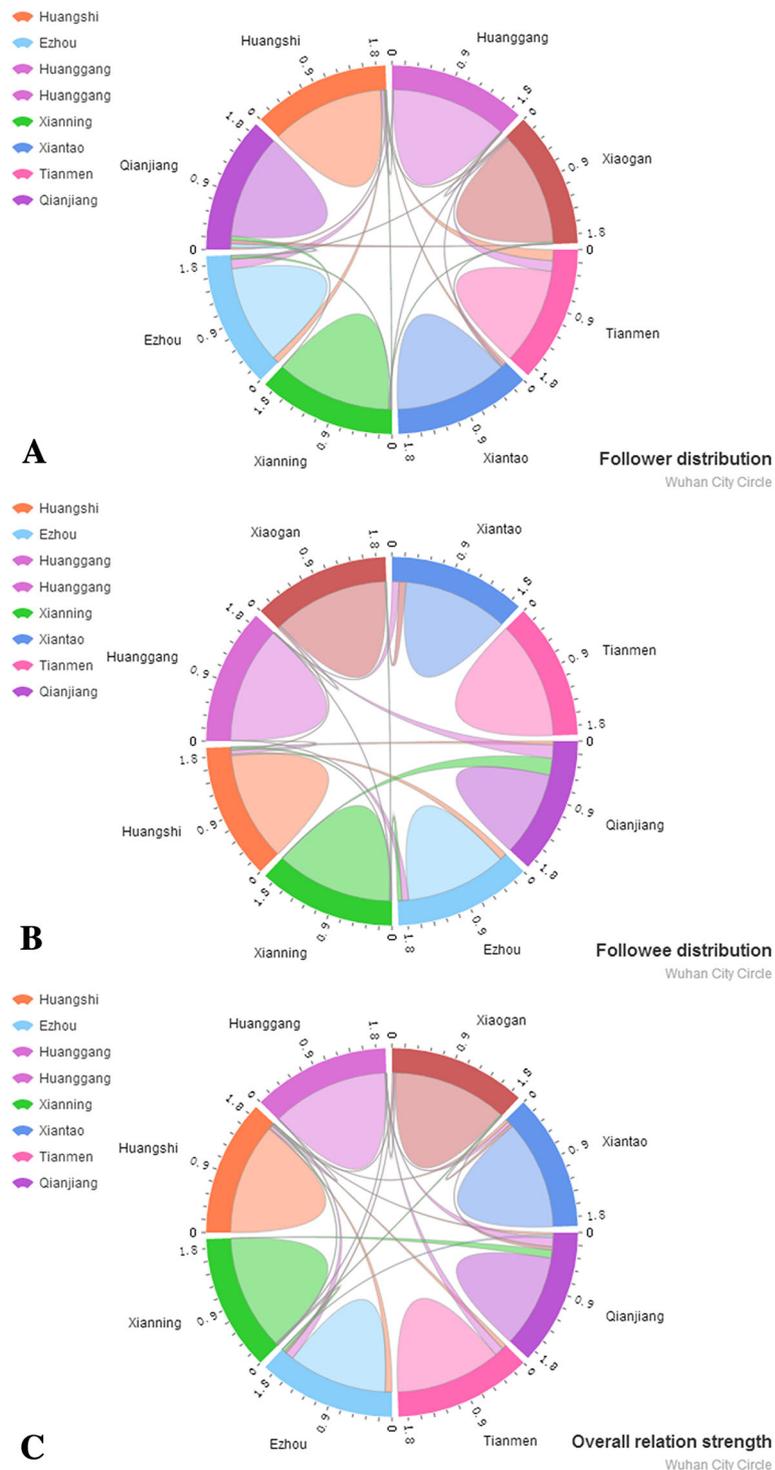
Fig. 10 Distributions of the relation strengths of eight cities (Excluding Wuhan). **a** Follower distribution. **b** Followee distribution. **c** Total distribution

the actual geographic space as the political, economic and cultural centre of Hubei Province.

- Our study indicated tiny communication among the other 8 cities along the periphery of WCC in the

virtual microblogging network space, with the exception of Wuhan; they closely resembled eight independent individuals. In particular, Huangshi lost the status of "the second centre city" of

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 13 of 14

WCC, and the distance between Xiantao and the centre city of the west wing increased. This finding is inconsistent with the reality in the actual geographical space, which indicates that the virtual network space resembled differed from the actual geographical space.

- The analysis of the WCC intercity information exchange characteristics indicated that the concept of WCC did not stand out in the virtual microblogging network space and that these cities frequently exchanged information with cities outside WCC.

- In an urban network, the degree of cohesion and outreach in a city node were significantly positively correlated with the economic development level of the city. A city with a high GDP rank frequently had a high degree of cohesion (and a low corresponding degree of outreach) and was more likely to engage in information exchange within the city and attract more attention from outside. Cities with a relatively low GDP were more inclined to engage in information exchange with cities with higher GDPs and tended to be more open and inclusive in the information space.

Additional file

Additional file 1: Table S1. Data statistics. (XLSX 9 kb)

Authors' contributions
SY and NZ and KW conceived and designed the experiments; SY and NZ and ZZ performed the experiments; NZ and KW and ZS and DS analyzed the data; SY and NZ wrote the paper. All authors read and approved the final manuscript.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]School of Resource and Environmental Sciences, Wuhan University, 129 Luoyu road, Wuhan 430079, China. [2]Ministry of Education, Key Laboratory of Geographic Information System (Wuhan University), Wuhan, China. [3]Shenzhen Research Center of Digital City Engineering, Shenzhen, China. [4]Research Center of Government Geographic Information System, Chinese Academy of Surveying and Mapping, Beijing, China.

References
1. Zhen F, Liu X, Liu H. Regional urban network influenced by information technology: new direction of urban studies. Hum Geogr. 2007;22:76–80.
2. Wu W, Cao Y, Liang S, Cao W. The accessibility pattern of railway passenger transport network in China. Geogr Res. 2009;28:1389–400.
3. Xue J. Hierarchical structure and distribution pattern of Chinese urban system based on aviation network. Geogr Res. 2008;27:23–33.
4. Zhen F, Wang B, Chen Y. China's city network characteristics based on social network space: an empirical analysis of Sina micro-blog. Acta Geograph Sin. 2012;67:1031–43.
5. Malecki EJ. The economic geography of the internet's infrastructure. Econ Geogr. 2002;78:399–424.
6. Rutherford J, Gillespie A, Richardson R. The territoriality of Pan-European telecommunications backbone networks. J Urban Technol. 2004;11:1–34.
7. Townsend AM. Network cities and the global structure of the internet. Am Behav Sci. 2001;44:1697–716.
8. Kwan MP. Mobile communications, social networks, and urban travel: hypertext as a new metaphor for conceptualizing spatial interaction. Prof Geogr. 2007;59:434–46.
9. Li Q, Chang X, Shaw S, Yan K, Yue Y, Chen B. Characteristics of micro-blog inter-city social interactions in China. J Shenzhen Univ Sci Eng. 2013;30:441–9.
10. Qin X, Zhen F, Xiong L, Zhu S. Methods in urban temporal and spatial behavior research in the big data era. Prog Geogr. 2013;32:1352–61.
11. Yu H, Shaw SL. Exploring potential human activities in physical and virtual spaces: a spatio-temporal GIS approach. Int J Geogr Inf Sci. 2008;22:409–30.
12. Chen Y, Zhen F, Wang B, Zou W. A study of internet information asymmetry relations among Chinese cities based on the micro-blog platform. Adv Earth Sci. 2012;27:1353–62.
13. Ding Z, Jia Y, Zhou B. Survey of data mining for micro-blogs. J Comp Res Dev. 2014;51:691–706.
14. Mausam RA, Etzioni O, Clark S. Open domain event extraction from twitter. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China. 2012. p. 1104–12.
15. Yang X, Ghoting A, Ruan Y, Parthasarathy S. A framework for summarizing and analyzing twitter feeds. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA. 2012. p. 370–8.
16. Bollen J, Mao H, Zeng X. Twitter mood predicts the stock market. J Comput Sci. 2011;2:1–8.
17. Thelwall M, Buckley K, Paltoglou G. Sentiment in twitter events. J Am Soc Inf Sci Technol. 2011;62:406–18.
18. Spina D, Meij E, de Rijke M, Oghina A, Bui MT, Breuss M. Identifying entity aspects in micro-blog posts. In: Proceedings of the 35th International ACM SIGIR Conference on Research and Development in Information Retrieval, Portland, OR, USA. 2012. p. 1089–90.
19. Yao J, Cui B, Xue Z, Liu Q. Provenance-based indexing support in micro-blog platforms. In: Proceedings of the 2012 IEEE 28th International Conference on Data Engineering (ICDE), Piscataway, NJ, USA. 2012. p. 558–69.
20. Recuero R, Araujo R, Zago G. How does social capital affect retweets? In: Proceedings of the 5th International AAAI Conference on Weblogs and Social Media, Menlo Park, CA, USA; 2011.
21. Yin D, Hong L, Davison BD. Structural link analysis and prediction in microblogs. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, New York, NY, USA. 2011. p. 1163–8.
22. Naaman M, Zhang AX, Brody S, Lotan G. On the study of diurnal urban routines on twitter. In: Proceedings of the 6th International AAAI Conference on Weblogs and Social Media, Palo Alto, CA, USA. 2012.
23. Xue Z, Yao J, Cui B. Temporal provenance discovery in micro-blog message streams (abstract only). In: Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data, Scottsdale, AZ, USA. 2012. p. 864-864.
24. Ding Z, Jia Y, Zhou B, Zhang J, Han Y, Yu C, et al. An influence strength measurement via time-aware probabilistic generative model for microblogs. In: 15th Asia-Pacific web conference on Web technologies and Applications, Sydney, Austrilian, 4-6 April 2013. Berlin: Springer; 2013. p. 372–83.

Ying *et al. Open Geospatial Data, Software and Standards* (2017) 2:2

Page 14 of 14

25. Steeg GV, Galstyan A. Information transfer in social media. In: Proceedings of the 21st International Conference on World Wide Web, New York, NY, USA. 2012. p. 509–18.

26. Li L, Yang L, Zhu H, Dai R. Explorative analysis of Wuhan intra-urban human mobility using social media check-in data. PLoS ONE. 2015;10:e0135286.

27. Wu L, Zhi Y, Sui Z, Liu Y. Intra-urban human mobility and activity transition: evidence from social media check-in data. PLoS ONE. 2014;9:e97010.

28. Wang B, Zhen F, Xi G, Qian Q, Wu C, Zhang H. A study of cybergeography based on micro-blog users' relationship: with a case of Sina micro-blog. Geogr Res. 2013;32:380–91.

29. Chen Y, Zhen F, Wang B, Zou W. Chinese city network structure in the cyberspace based on social network analysis. Econ Geogr. 2013;33:56–63.

30. Chang X, Yue Y, Li Q, Chen B, Xiao S, Tu W. Extracting the geographic backbone of location-based social network. Geo Spat Inf Sci. 2014;39:706–10.

31. Cener SWD. Sina Weibo user report. 2016. http://tech.sina.com.cn/2016-12-15/doc-ifxytqax6071840.shtml. Accessed 20 Dec. 2016.

32. Wang J, Zhu K, Wang B. Survey on microblog research based on information data analysis. J Comput Appl. 2012;7:65.

33. Golder SA, Yardi S. Structural predictors of tie formation in twitter: transitivity and mutuality. In: Proceedings of the 2010 IEEE Second International Conference on Social Computing, 2010. Piscataway: IEEE Computer Society; 2010. p. 88–95.

34. Boyd D, Golder S, Lotan G. Tweet, tweet, retweet: conversational aspects of retweeting on twitter. In: Proceedings of the 43rd Hawaii International Conference on System Sciences (HICSS), 5-8 January 2010. Honolulu: IEEE; 2010. p. 1–10.

35. Granovetter MS. The strength of weak ties. Am J Sociol. 1973;78:1360–80.

36. He J, Shi X, Gong J. Yu Yan. Modeling the spatial expansion of urban agglomeration considering their spatial interaction: a case study of Wuhan Metropolitan Area. Geo Spat Inf Sci. 2016;41:462–7.

37. Liu C, Yu R. Spatial Accessibility of Road Network in Wuhan Metropolitan Area Based on Spatial Syntax. J Geogr Inf Syst. 2012;04:128–35.

38. He J, Li C, Liu Y, Yu Y. A field source-strength method for interaction scenario in network space of Metropolitan. Acta Geodaetica CartographSin. 2015;44:805–12.

39. Yuan F, Feng J, Fu Q, Cao X. A method to reduce the impact of zombie fans in micro-blog. J Am Soc Inf Sci. 2012;5:70–5.

40. Lian J, Zhou X, Cao W, Liu Y. Sina microblog data retrieval. Tsinghua Sci Technol. 2011;51:1300–5.

41. Zhou Z, Zhang H, Xie J. Data crawler for Sina Weibo based on Python. J Comput Appl. 2014;34:3131–4.