To Cite:

**NETWORK ANALYSIS OF NAVIGATION PATHS OF TOURISTS' TRIP PLANNING AND POWER STRUCTURE OF THE ONLINE TOURISM IN CHINA**

Junyi Li, Ph.D.
Assistant Professor
College of Tourism and Environment
Shaanxi Normal University, Xi'an, China
Telephone: 86-133-79263086
Fax: 86-29-85310495
Email: lijunyi9@snnu.edu.cn

Min Yang, Ph.D.
School of Business
Xi'an University of Finance and Economics, Xi'an, China
Telephone: 86-153-39058773
Fax: 86-29-81556357
E-mail: amandamin@163.com

Bing Pan, Ph.D.*
Associate Professor
Department of Hospitality and Tourism Management
School of Business
College of Charleston, Charleston, SC 29424-001, USA
Telephone: 1-843-953-2025
Fax: 1-843-953-5697
E-mail: bingpan@gmail.com

* Corresponding Author

**Acknowledgments**: The project was jointly funded by National Natural Science Foundation Projects No. 41001077 and No. 41428101, China Postdoctoral Science Foundation Sponsored Project No. 2012T50794, China Postdoctoral Science
Abstract

In order to investigate the power structure of the online tourism domain in China, this study adopted a lab experiment setting and a network analysis method to examine the navigational paths of vacation planning behavior of Chinese college students. The analyses included the websites involved in the information search, network statistics of navigational paths, and clickthrough patterns between different types of websites. The navigational path network is a small-world network with shorter average path lengths and shorter network diameter than a website hyperlink network. The tourists generally followed a path of starting web page, general search engines, OTA, to specific business website. The clickthrough pattern indicated that search engines and Online Travel Agencies (OTAs) were major intermediaries and thus exerted a higher power in the online tourism domain in China.

Keywords: online tourism; information search; tourism websites; website clickthrough; social network analysis; complex networks
NETWORK ANALYSIS OF NAVIGATION PATHS OF TOURISTS' TRIP PLANNING
AND POWER STRUCTURE OF ONLINE TOURISM DOMAIN IN CHINA

INTRODUCTION

With the fast development of the Internet and e-commerce in tourism, the way tourists access information has also undergone dramatic changes: online information has become the most important information source for tourists, including those in China (Lehto, Kim and Morrison 2006; Li and Yang 2010; Zhang, Liang, Shi and Zhong 2007). However, the development of tourism e-commerce in China lags behind that of other developed countries. The proportion of online transactions is relatively small (China Tourism Academy 2011), as the number of online tourism searchers far exceeds that of online bookers (Li and Yang 2010). The investigation of online tourist information search behavior is crucial for understanding the consumer behavior of tourists, and thus, for the management of tourism e-commerce in China (Raitz and Dakhil 1989; Schul and Crompton 1983; Snepenger, Meged, Snelling and Worrall 1990).

Researchers have conducted many studies investigating tourist information search behavior and the problems during the process of the search. Fodness and Murray (1997) deem tourism information search as a process of dynamic development wherein individuals adopt various types and amounts of information sources in order to facilitate travel planning. Another study by Lehto, Kim and Morrison (2006) investigated the involvement, consumer cognition process, and knowledge characteristics of searchers during tourist information search. Scholars also found various priority levels for different aspects of a trip, starting from
transportation, eating, accommodation, to shopping, entertainment, and attractions, in that sequence (Li and Yang 2010). Pan and Fesenmaier (2006) conducted a lab study on online tourist information search and vacation planning, and revealed that tourists often browse information hubs with lots of links, e.g. websites containing a large number of official tourism websites and hotel websites. Their study also found that large discrepancy exists between the semantics of tourists versus that of information providers. Some recent studies on tourist information search also concentrate on the search behavior within a website (Hodkinson, Kiel and McColl-Kennedy 2000; Srivastana, Cooley, Deshpande and Tan 2000), while few studies exist on the clickthroughs between them.

Researchers have investigated the so-called online tourism domain, the aggregated information space related to tourism. Xiang, Wober, and Fesenmaier (2008) found that in the online tourism domain regarding destinations, online a small fraction of websites are visible on search engines. Some studies investigated the power-structure of information space through the lenses of visible information of search engines. There are huge discrepancy between the semantics of online tourism and the tourists’ semantic when they search for information (Xiang, Gretzel, and Fesenmaier 2009). However, tourists browse online travel information through continuous clicks among various sites and pages, which goes beyond search engines.

The studies by Pan and Fesenmaier (2006) and Zhang (2010) explored clickthroughs between websites; nonetheless, they did not include in-depth analysis of tourists’ navigation paths. The sequence of navigation by tourists is a directed graph, which can be analyzed using statistical methods (Wasserman & Faust 1994). Formal network analysis could reveal important structural information about tourists'
complete navigation paths. This will go beyond the search engine result pages or the navigation path inside one website and thus give a more accurate picture on who the power players are in the online tourism information space. What are the websites which the tourists encounter during their whole search process? What are the most important websites in their navigation? This study adopts network analysis methods to formally investigate the characteristics of the complete navigation paths of Chinese tourists, and the clickthrough networks they form when planning their trips. The results reveal the detailed power-structure of China’s online tourism domain.

LITERATURE REVIEW

This section reviews relevant studies in tourist information search, online tourism domain, and network analysis. Gaps in research on tourist information search and trip planning are discussed.

Tourist Information Search

In many ways, tourists search for information differently from other consumers, in that a tourist’s search involves many different aspects of a trip (Kim, Lehto and Morrison 2007), and it has been widely studied in the field of tourism (Fodness and Murray 1997; Chen and Gursoy 2000; Gursoy 2003; Gursoy and McCleary 2004). The informational needs of tourists include functional, hedonic, innovational, aesthetic, and sign needs (Vogt and Fesenmaier 1998). Compared to the traditional sources of information, online information not only saves users’ time, cost and energy, but also provides more reliable tourism information (O’Connor 1999). Jepson’s 2007 study finds that the low cost of an online search is one of the important reasons people use the web to find information. Pan and Fesenmaier
(2006) found that tourists’ planning can be deconstructed into a series of episodes and chapters reflecting the specific problem being addressed. Tourists often browse the information hubs, which contain a large number of official tourist activity websites and accommodation websites (Pan and Fesenmaier 2006). There are large discrepancy between the languages the tourists use versus that of tourist information providers. However, no in-depth analysis has been conducted on those navigation paths.

Research on Online Tourism Domain

A few studies have investigated online information through information supplier’s perspective and compared it with tourists’ perspective. Xiang, Wober, and Fesenmaier (2008) defined online tourism domain as the aggregate of information entities and technological elements related to travel. It includes texts, images, video/audio, and hypertext on the Internet. The researchers fed travel-related queries in Google and found that the visible travel information through search engines is dominated by a small number of websites. Further, there are huge discrepancy between the semantics of online tourism and the tourists’ semantic when they search for information (Xiang, Gretzel, and Fesenmaier 2009).

Using a similar methodology of data mining on retrieved search engine results, Xiang and Greztel (2010) discovered that social media sites are embedded with tourists’ search results on Google and they are highly visible. Many times, search engines direct the travel information searchers to those social media sites. These studies investigated the power-structure of information space through the lenses of
visible information of search engines. This is reflected from the keywords the two parties are using and the relative importance of those keywords.

In general, these studies investigated online information through the lenses of search engines. However, tourists’ online navigation goes beyond search engines, including travel portals, destination websites, Online Travel Agencies, etc. An in-depth analysis of tourists’ online path for searching and navigation could afford more insights in the power-structure of online tourism. What are the websites which are accessed during the process? What websites are the central sites in connecting tourists with online information? A web users’ clicks can be viewed as a directed graph, which can be subject to network analysis methods (Wasserman and Faust 1994). This method could help reveal the central nodes in tourists’ navigation paths.

**Network Analysis Methods**

Garton, Haythornthwaite and Wellman (1997) and Jackson (1997) suggested that social network analysis method, one of the complex network methods, is a useful approach to studying the representation and interpretation of the online communication structure and computer-mediated communication (CMC) processes. With only a few empirical studies, the research on tourism hyperlink networks is still in its early stages, and focuses on the hyperlink structure in online tourism space. Baggio, Corigliano and Tallinucci (2007) explored the topological characteristics of a destination network. Luciano da Fontoura Costa and Rodolfo Baggio (2009) reported a complex network analysis of a tourism destination website network based on Elba, Italy, including the characterization of its structure. The study revealed the network
structural characteristics, including the number of paths and distinct lengths between pairs of nodes, as well as the number of reachable companies, and its dynamic features, including transition probabilities and the inward/outward activations and accessibilities. This showed that the type and size of the companies correlated strongly with their respective activations and accessibilities. The study also showed that the Elba tourism network is largely fragmented and heterogeneous, so that it could benefit from increased integration. In addition, Miguéns and Corfu (2008) analyzed the online connectivity of tourist attractions in Lisbon and revealed the disassortativity in tourism attractions’ online network. Piazzi, Baggio, Neidhardt and Werthner (2012) compared the web space structure of two destinations and confirmed the usefulness of quantitative network measures in characterizing the online collaboration of tourism stakeholders in destinations. Social network analysis has also been used to investigate the spatial movement of international tourists in China (Leung, Wang, Wu, Bai, Stahura and Xie 2012). However, very few studies have investigated the navigational paths through network structure.

As one exception, Zhang (2010) modeled Chinese consumers’ cross-website transfer behavior. He applied Zipf’s law and complex network theory and compared the way two types of information searchers, consumers involved in a specific task and average web users in an everyday life scenario, would allocate search efforts and transfer among websites. He found that their cross-site behaviors were similar: they allocated their efforts among websites according to Zipf’s law. However, the topology characteristics of their transferring networks were both similar and different. For
example, their small world nature, high clustering coefficient, disassortative mixing and poor invulnerability, were similar to general web user networks, but other measurements were different, such as not following power-law degree distribution, and having circular and self-similar community structure.

In general, past studies on tourist information search behavior follow a consumer research perspective. A few researchers investigated the distribution of online tourism domain and the power relationships of online information, but only through the lenses of search engines. Zhang (2010)’s study focused on general consumer information seeking, rather than travel information search. Pan and Fesenmaier (2006) did report general tourist information search behavior through graph representation, but did not adopt a formal network analysis method. No studies have investigated Chinese tourists' navigational paths through a network analysis method. A network analysis could reveal the structural characteristics of the paths; for example, the size of the navigational paths, and the crucial players in the graph. This can reveal the detailed power structure of online tourism domain for Chinese tourist. The current study adopted a lab experiment setting and used network analysis methods to formally study the navigation paths of Chinese tourists when they plan a trip. The following details the research methodology.

2. Research Methodology

In order to understand the characteristics of Chinese tourists' trip planning behavior and the power structure of the online tourism domain in China, a lab experiment was conducted. This section details the data collection process and
2.2 Data collection and processing

The subjects were recruited with a convenient sampling method. Since we are investigating the innate characteristics of online tourist information search behavior, they were considered an appropriate population for this study. The subjects were mainly recruited from four universities in a large city in the west of China. The researchers sent an invitation to students in several classes in those universities, and they were informed that they could invite their classmates or friends to join. The final participants also included students from other universities in China. Prior to the experiment, the researcher copied the experimental requirement and instruction document, screen recording software, and survey forms to form a software package, emailed them to the students who agreed to participate, and asked them to complete the experiment during their discretionary time, with no time or location constraints. The subjects were asked to email back the final dataset in the form of a package.

This study used a piece of screen capturing software for PCs (Screen Recorder V7.5, TechSmith, 2013) to record the online tourist information search process of each student. The task was to plan a seven-day sightseeing visit to Beijing with 2,000 RMB (about $310 USD in 2011) as the travel budget. The subjects were asked to search for information online and formulate a detailed one-week sightseeing plan. Before the formal experiment, the subjects needed to read about the purpose and significance of the study, familiarize themselves with the installation and operation of the screen capturing software, read through the experiment requirements, and fill out a pre-experiment questionnaire. The questionnaire and screen-recording data had unique identification numbers. No limit was set on the amount of time the subjects
could take to plan their trips.

From September to November in 2011, 166 experimental sessions were completed and 136 valid experimental video data points were obtained by excluding data with extremely short search time (less than 10 minutes) or with incomplete information.

Though the data were not collected in a strict lab environment, many conditions validated the study method. We only used student population; the subjects can only use PC as the platform as instructed, since the screen capturing software is only provided for a PC platform; we also eliminated 30 videos of less than 10 minutes from a total of 166 videos; the recorded videos also allowed us to see if anything suspicious happening. The 136 videos passed face validity test. In general, we made the best effort in controlling certain variables. Thus, our experiment conditions are not significantly different from experiments in a lab setting, as done in Pan and Fesenmaier (2006).

2.3 Analysis Methods

We formalized the information searchers' navigational paths as directed networks, and describe here the relevant network analysis methods. A screen recording experiment was conducted to record the website clickthrough behaviors of the subjects in order to obtain the directed and weighted clickthrough matrix about the website visit behavior, and the network characteristics of visitors’ online information search were calculated.

When a tourist searched for information online, each site visited is a node (N). The website clickthrough path from \( N_i \) to \( N_j \) is recorded as a directed edge \( (N_i, N_j) \). The weight of directed edge \( F_{ij} \) is the number of times a user clicked from website \( N_i \) to website \( N_j \).
Node degree and centrality: the degree of nodes refers to the number of edges of one node that connected with others in the network. In a complex network, \( k_i \), the degree of node \( N_i \), is defined as the number of other nodes to which it directly connected. For a directed graph, the degree of a node is divided into “input degree” and “output degree”. The input degree is defined as the number of nodes pointing at node \( N_i \), while output degree is the number of nodes pointed to by node \( N_i \). The sum of output degree and input degree is the total degree of node \( N_i \) (Wasserman and Faust 1994). In a social network, the degree represents the influence and importance of certain individuals in the network. The greater the individual degree is, the greater its influence in the entire network and the greater its role throughout the entire organization, and vice versa. The researchers can also calculate an output degree/input degree ratio, that is, the ratio of input and output degree of each node. Under normal circumstances, nodes with a higher output degree are information hubs, since they lead to a large amount of other web links. Nodes with a higher input degree are nodes with high confidence, since they are the websites many others lead to. Therefore, if the ratio of output to input degree is greater than 1, they are the information originating centers; if less than 1, they are more likely to be the authoritative websites. Researchers can also calculate the centrality of the entire network and of each individual node. Centrality measures how “important” or “prominent” a node is in a network (Wasserman and Faust 1994). Centrality is calculated as the percentage of the number of degrees of an individual node compared to maximum possible degrees. Furthermore, group centrality measures how different centrality values are in a
network (Borgatti, Everett & Freeman, 1999).

*Average path length:* The distance among nodes refers to the minimum number of edges from one node to another, wherein the maximum distance among all node pairs is network diameter. The average path length (APL) refers to the average distance among all nodes in the complex network (Wasserman and Faust 1994). Average network path length and diameter are indicators of transmission performance and efficiency of the network.

*Clustering coefficient:* The clustering coefficient, also known as the convergence coefficient, is an indicator of the level of collectiveness of the network. For instance, the clustering coefficient $C_i$ of node $i$ describes its connecting relation with other conjoint nodes in the complex network; that is, the proportion of existing number of edges in relation to the possible maximum number of edges between this node and its directly adjacent nodes.

*Network elasticity:* Network elasticity was measured by first deleting different nodes in the network and then measuring the resulting connectivity of the network and the change in network diameter. Random deletion of network nodes affects the structure and indicators of the entire network.

**Results**

This section details the demographic information of the subjects, general information about the subjects' trip planning behavior, and the results of the network analysis. Detailed power structure of online tourism domain was revealed.
1. Demographics

Among all the subjects for this study, 69.1% were females. The subjects were mainly sophomore and junior college students, (38.9% and 37.4% respectively), followed by seniors (22.9%). Half of the students had a major in humanities and social science (49.6%), and the rest had a major in sciences and engineering (50.4%). The average amount of their lifetime during which they have been using the Internet was 6.9 years, with the shortest time of 1 year and the longest time of 14 years. A large proportion of students had between 3-7 years of web use history (41.9%). Around 73.3% students had not yet been to Beijing, 13.7% had been to Beijing once, and 13% had been there two or more times.

2. General Trip Planning Behavior

In total, 136 subjects visited 183 different websites. The subjects browsed and searched information for an average of 32 minutes and 18 seconds; the minimum criterion for inclusion for the analysis was 10 minutes, and the longest time was 102 minutes. About 42.6% of the subjects spent 20 to 60 minutes planning their visit. The average search time was 38.6 minutes for female students, significantly higher than 26.7 minutes for males. As to the locations where the trip planning exercise took place, 56.6% of the subjects used dormitories, followed by 17.9% in Internet bars. The proportions of other places (homes, laboratories and computer rooms) were small (Table 1).

-------- Insert Table 1 here  -----------
3. Network analysis of tourism information search

After collecting the videos and surveys, the researchers analyzed the content of the screen capture videos. The subjects' clickstreams were coded as a stream of websites visited. The researchers then conducted structural network analysis on the navigation paths of the subjects, including the most visited websites, the degree distribution of the network, average path length, network diameter, clustering coefficient, and network elasticity. The network characteristics of the navigational paths reflected the overall structure of the network, and the following analysis further explores detailed behavior patterns.

3.1 General Search Patterns

The researchers analyzed the top 21 most visited websites, which were visited by 5 or more subjects. The websites’ clickthrough matrix was constructed. The results showed that among the total clickthroughs for 21 websites, Baidu accounted for 40.9%, much higher than those of others, such as Qunar (8.9%), Ctrip (7.3%) and Cncn (6.9%). Qunar, Ctrip and Cncn are popular OTA (Online Travel Agency) sites in China. The most frequent clickthrough was from Baidu to Qunar, which accounted for 16.8% of Baidu’s total clickthroughs. The most searched-for content was train tickets; Qunar appeared at the top position of the result page, and the train schedule information was displayed with an enlarged font. The clickthrough number from Qunar to Baidu reached 33, which accounted for 16.3% of the entire clickthrough. Among all websites, the access frequency of Baidu ranked first, followed by Qunar. Between-website clickthroughs mainly happened between search engines (Baidu and
Sogou, the two most popular search engines in China) and OTA websites (Qunar and Cncn). These sites occupied important positions, and the deletion of them would greatly affect the structure of the entire network. The main patterns of Chinese tourists' online navigation paths were from starting website → general search engines → OTA sites → specific business websites. Thus, the first intermediary sites were generally search engine sites such as Baidu, Google, Soso, Sogou, etc. The search engine most frequently used by domestic tourists was Baidu. The second level intermediate websites were generally OTA websites such as Qunar, Ctrip, 17U and Cncn.

3.2 Analysis of degree distribution

The nodes with high degrees were Baidu (the most popular search engine in China), Sogou (search engine), Cncn (OTA), and Google (search engine), as well as Qunar (OTA), Uzai (OTA), Ctrip (OTA), 17U (OTA), lvmama (DIY travel), and hao123 (an online directory) (Table 2). The core of the hub was search engine website, while “Baidu” was the most important node in the entire information search, followed by OTAs, such as Qunar, Ctrip, and 17U.

----- Insert Table 2 here -------

By comparing the output and input degrees of the 10 most important nodes, the researchers found that websites with ratios greater than 1 or equivalent to 1 included Yododo, Sina, Sohu, Sogou, Google, Baidu, hao123 and hao360 (Table 3). All of these sites were search engines, site navigation and news portals, except Yododo. The ratios of input-output degree of Baidu, Sohu and Sina were close or equivalent to one,
which indicated that these sites were considered to be website information authorities. In contrast, all OTA sites had an input-output degree of less than one.

----- Insert Table 3 here ---------

3.3 Analysis of average path length and network diameter

The average path length was calculated as 2.6 \((\log n / \log \log n = 6.38, n = 341,183)\) by analyzing the website clickthrough matrix of the experimental data via PAJEK software (Batagelj & Mrvar, 2004), the diameter of the network was 6.0, and there were 23,570 reachable pairs and 9736 unreachable pairs, each taking up 70.8% and 29.3% of the total node pairs \((n \times (n-1) = 33,306)\), respectively. The result showed that in the navigation path network, the nodes (websites) were not distant from each other. That is, the network of tourists' online navigational paths was in-line with the characteristics of a small-world network. However, compared to the tourism network link of Elba Island, the paths and diameter were both shorter: Elba Island was 4.5 in average path and 11 in diameter (Costa and Baggio 2009). These were also shorter than the average network path length (3.97) and network diameter (44) of whole website linkage according to ComScore's data reported by Zhang in 2010 (Srivastana, Cooley, Deshpande & Tan 2000). These indicated web navigation paths were an even smaller network, due to their confinement within a small part of online tourism domain.

3.4 Analysis of clustering coefficient

The global clustering coefficient (GC) of the network and the close neighbor clustering coefficient (LC) were calculated as 0.03 and 0.36, respectively, by using network software PAJEK. They were relatively higher values than the network structure of the whole web (the GC and LC of website linkage network are 0.008 and
0.11, respectively, according to ComScore [Srivastava et al. 2000]). These indicated the navigational path networks were even better connected. The websites were close to each other due to the small network size.

3.5 Analysis of network elasticity

Table 4 shows that random deletion of nodes in the network had limited impact on the network diameter. Possibly due to its small size (183 nodes), the deletion of key nodes had great impact, for instance, with the deletion of the node of Baidu. In that case, the network diameter changed from 6 to 9, the average path length increased from 2.56 to 3.86, and unreachable nodes amounted to 77.3%. Thus, key nodes had large impact on the online tourism network structure in the network, and these nodes were crucial for online navigation. While deleting 2-10 ordinary nodes, the average path length changed to from 2.55 to 2.54, while the network diameter remained at 6.

4. Conclusions and Discussions

The study used a lab experiment to investigate the navigation paths of Chinese trip planners. Under the framework of complex network theory, it analyzed the complex network structural features including degree distribution, node centricity, network diameter and average path length, clustering coefficient and network elasticity. The results showed that the Chinese tourists' navigational paths formed a small network. Its average length paths and network diameter were all smaller than that of website hyperlink structure of the whole web or a part of the web on a tourist destination. This indicated that the navigational paths were always confined by a
small amount of websites inside the web network. It also indicated that tourists might generally follow a small limited space within the destination website network, which leads to a smaller diameter.

Based on the network analysis of navigational paths of Chinese tourists, a few patterns emerged from the analysis. By comparing the output degree versus input degree, output degrees of general search engines portals were found to be significantly greater than those of input degrees, while for specific attraction or tourism business websites the input degree was significantly greater than that of the output degree.

Chinese trip planners usually went from search engines to OTAs as the main information search route. General search engines were the hubs for the Chinese tourists' navigational paths. The secondary level of hubs was OTA websites (such as Qunar and Cncn). These sites had high input and output degrees and occupied important positions in the whole clickstream. The deletion of these nodes greatly affected structural characteristics of the entire navigational path network.

This study reveals the power structure of online tourism domain in China. Specifically, search engines and OTAs have tremendous influence on where trip planners go online. These two types of websites dominated the Chinese tourists’ online navigation paths and occupied central positions. Due to their expertise in online marketing, OTAs are highly visible search engine result pages when tourists search for information. The ranking of OTAs on general search engines contributes to their dominant power. This confirms previous studies in that small and media
businesses can hardly be visible in search engine space (Xiang, Wober, Fesenmaier, 2008). Thus, the individual businesses have limited choice but to work with them due to their visibility on search engines and as a consequence, prominent positions in tourists' online navigational path network. This gives OTAs great power in dictating the commissions and relationships with hotels and attractions. This contributed to the love-hate relationship between OTA sites and individual hotel or travel websites, both in China and the developed countries. This calls for more training for advanced online marketing techniques such as search engine marketing for small to medium tourism and hospitality businesses. However, with a limited budget, it might be hard for those businesses to make such investment.

The limitation of this study lies in its experimental nature. The college students were asked to plan an imagined trip to Beijing with a set budget. The generalization of the results might, therefore, be limited. Future studies could expand the samples by capturing the navigational patterns of population-balanced Chinese tourists.
References


### Table 1. Demographics of Subjects (%)

<table>
<thead>
<tr>
<th>Items</th>
<th>Options</th>
<th>%</th>
<th>Items</th>
<th>Options</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>male</td>
<td>30.9</td>
<td>Major</td>
<td>Arts</td>
<td>49.6</td>
</tr>
<tr>
<td></td>
<td>female</td>
<td>69.1</td>
<td></td>
<td>Science and engineering</td>
<td>50.4</td>
</tr>
<tr>
<td>Freshman</td>
<td>0.8</td>
<td></td>
<td>Length of vacation</td>
<td>Less than 15 min.</td>
<td>30.15</td>
</tr>
<tr>
<td>Grade</td>
<td>Sophomore</td>
<td>38.9</td>
<td>planning time</td>
<td>15-30 min.</td>
<td>27.21</td>
</tr>
<tr>
<td></td>
<td>Junior</td>
<td>37.4</td>
<td></td>
<td>30-60 min.</td>
<td>28.68</td>
</tr>
<tr>
<td></td>
<td>Senior</td>
<td>22.9</td>
<td></td>
<td>60-120 min.</td>
<td>13.97</td>
</tr>
<tr>
<td>Online history</td>
<td>&lt; 3 years</td>
<td>11.3</td>
<td></td>
<td>Dorm</td>
<td>56.6</td>
</tr>
<tr>
<td></td>
<td>3-7</td>
<td>41.9</td>
<td></td>
<td>Internet bar</td>
<td>17.9</td>
</tr>
<tr>
<td></td>
<td>7-10</td>
<td>31.5</td>
<td>Locations where</td>
<td>Library</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>&gt; 10 years</td>
<td>15.3</td>
<td>subjects accessed the Internet</td>
<td>Home</td>
<td>9.7</td>
</tr>
<tr>
<td>Times visited</td>
<td>Never</td>
<td>73.3</td>
<td></td>
<td>Laboratory</td>
<td>6.2</td>
</tr>
<tr>
<td></td>
<td>Once</td>
<td>13.7</td>
<td></td>
<td>Cell phone</td>
<td>2.1</td>
</tr>
<tr>
<td>Beijing</td>
<td>Twice</td>
<td>8.4</td>
<td></td>
<td>School computer room</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td>Over 3 times</td>
<td>4.6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Websites</td>
<td>Degree centrality</td>
<td>Node degree</td>
<td>Type</td>
<td>Websites</td>
<td>Degree centrality</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------------</td>
<td>-------------</td>
<td>--------------</td>
<td>-----------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Baidu</td>
<td>0.577</td>
<td>105</td>
<td>Search Engine</td>
<td>Lotour</td>
<td>0.027</td>
</tr>
<tr>
<td>Sogou</td>
<td>0.115</td>
<td>21</td>
<td>Search Engine</td>
<td>Quxun</td>
<td>0.027</td>
</tr>
<tr>
<td>Cncn</td>
<td>0.088</td>
<td>16</td>
<td>OTA</td>
<td>Douban</td>
<td>0.027</td>
</tr>
<tr>
<td>Google</td>
<td>0.066</td>
<td>12</td>
<td>Search Engine</td>
<td>Map bar</td>
<td>0.027</td>
</tr>
<tr>
<td>Qunar</td>
<td>0.055</td>
<td>10</td>
<td>OTA</td>
<td>Yododo</td>
<td>0.022</td>
</tr>
<tr>
<td>Uzai</td>
<td>0.038</td>
<td>7</td>
<td>OTA</td>
<td>Elong</td>
<td>0.022</td>
</tr>
<tr>
<td>Ctrip</td>
<td>0.038</td>
<td>7</td>
<td>OTA</td>
<td>Tao123</td>
<td>0.022</td>
</tr>
<tr>
<td>17U</td>
<td>0.038</td>
<td>7</td>
<td>OTA</td>
<td>Soso</td>
<td>0.022</td>
</tr>
<tr>
<td>Lvmama</td>
<td>0.038</td>
<td>7</td>
<td>DIY tour</td>
<td>Dreams-tourism</td>
<td>0.022</td>
</tr>
<tr>
<td>Hao123 website navigation</td>
<td>0.038</td>
<td>7</td>
<td>Online Directory</td>
<td>Jinpai365</td>
<td>0.022</td>
</tr>
<tr>
<td>nettvl</td>
<td>0.033</td>
<td>6</td>
<td>OTA</td>
<td>Dianping</td>
<td>0.022</td>
</tr>
<tr>
<td>Tuniu</td>
<td>0.027</td>
<td>5</td>
<td>OTA</td>
<td>360 website navigation</td>
<td>0.022</td>
</tr>
<tr>
<td>Lvping</td>
<td>0.027</td>
<td>5</td>
<td>Social Media</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 3. The input degree-output degree results of the first 21 websites

<table>
<thead>
<tr>
<th>Websites</th>
<th>Yododo</th>
<th>Elong</th>
<th>Sina</th>
<th>CNCN</th>
<th>Ctrip</th>
<th>Tuniu</th>
<th>17U</th>
</tr>
</thead>
<tbody>
<tr>
<td>Output degree</td>
<td>1.14</td>
<td>0.80</td>
<td>1.00</td>
<td>0.68</td>
<td>0.83</td>
<td>0.80</td>
<td>0.75</td>
</tr>
<tr>
<td>input degree ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Websites</td>
<td>Sohu</td>
<td>Sogou</td>
<td>Qunar</td>
<td>Internet tourism</td>
<td>Lypyng</td>
<td>Lvmama</td>
<td>Lotour</td>
</tr>
<tr>
<td>Output degree</td>
<td>1.00</td>
<td>2.00</td>
<td>0.91</td>
<td>0.91</td>
<td>0.83</td>
<td>0.80</td>
<td>0.89</td>
</tr>
<tr>
<td>input degree ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Website</td>
<td>Kuxun</td>
<td>Google</td>
<td>Doubain</td>
<td>Daodao</td>
<td>Baidu</td>
<td>hao12</td>
<td>hao360</td>
</tr>
<tr>
<td>Output degree</td>
<td>0.96</td>
<td>1.18</td>
<td>1.00</td>
<td>0.50</td>
<td>1.02</td>
<td>3.43</td>
<td>3.75</td>
</tr>
<tr>
<td>input degree ratio</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4. Damage resistance ability of online tourist information search network

<table>
<thead>
<tr>
<th>Node</th>
<th>Average path length</th>
<th>Network diameter</th>
<th>Quantity of all nodes</th>
<th>Reachable pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>183</td>
<td>2.565</td>
<td>6</td>
<td>33306</td>
<td>23570</td>
</tr>
<tr>
<td>182</td>
<td>3.865</td>
<td>9</td>
<td>32942</td>
<td>7484</td>
</tr>
<tr>
<td>181</td>
<td>2.554</td>
<td>6</td>
<td>32580</td>
<td>22234</td>
</tr>
<tr>
<td>179</td>
<td>2.542</td>
<td>6</td>
<td>31862</td>
<td>20914</td>
</tr>
<tr>
<td>177</td>
<td>2.549</td>
<td>6</td>
<td>31152</td>
<td>19610</td>
</tr>
<tr>
<td>175</td>
<td>2.557</td>
<td>6</td>
<td>30450</td>
<td>18322</td>
</tr>
<tr>
<td>173</td>
<td>2.539</td>
<td>6</td>
<td>29756</td>
<td>15918</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Node</th>
<th>Proportion of reachable pairs</th>
<th>Unreachable pairs</th>
<th>Proportion of unreachable pairs</th>
<th>Deleted nodes</th>
</tr>
</thead>
<tbody>
<tr>
<td>183</td>
<td>70.77%</td>
<td>9736</td>
<td>29.23%</td>
<td></td>
</tr>
<tr>
<td>182</td>
<td>22.72%</td>
<td>25458</td>
<td>77.28%</td>
<td>Key node No.103</td>
</tr>
<tr>
<td>181</td>
<td>68.24%</td>
<td>10346</td>
<td>31.76%</td>
<td>2 ordinary nodes</td>
</tr>
<tr>
<td>179</td>
<td>65.64%</td>
<td>10948</td>
<td>34.36%</td>
<td>4 ordinary nodes</td>
</tr>
<tr>
<td>177</td>
<td>62.95%</td>
<td>11542</td>
<td>37.05%</td>
<td>6 ordinary nodes</td>
</tr>
<tr>
<td>175</td>
<td>60.17%</td>
<td>12128</td>
<td>39.83%</td>
<td>8 ordinary nodes</td>
</tr>
<tr>
<td>173</td>
<td>53.50%</td>
<td>13838</td>
<td>46.50%</td>
<td>10 ordinary nodes</td>
</tr>
</tbody>
</table>