An index model for measuring microblog users’ influence*

Fuyong YUAN¹, Jing FENG¹† & Qianqian FU¹,2

¹ College of Information Science and Engineering, Yanshan University, Qinhuangdao 066004, China
² Department of Economics, Qinhuangdao Institute of Technology, Qinhuangdao 066100, China

Abstract

Purpose: The paper aims to build an index model for measuring microblog users’ influence by taking microbloggers of Sina Weibo as a research sample.

Design/methodology/approach: Our user influence index model emphasizes link analysis and user activities in the microblogging network. We conduct experiments to investigate the performance of our model by using data crawled from Sina Weibo.

Findings: User influence is correlated to the attention that a user has received from his/her audience, the user’s activities and his/her tweets’ influence. Experimental results show that our model can reflect microbloggers’ influence in a more reasonable way.

Research limitations: More factors need to be considered to identify different influential users at different time periods.

Practical implications: The results of the study provide us with insights both into the way to measure microblog users’ influence and to rank users based on their influence.

Originality/value: By combining link analysis and user activities, this index model can reduce the impact of dummy follower accounts on user influence, reflecting a user’s real influence in the microblog system.

Keywords User influence; Audience’s attention; User activity; Tweets’ influence

1 Introduction

An increasingly growing number of Internet users has been accustomed to browsing information and expressing feelings in the virtual community with the prevalence of microblogs¹. This has changed the microblogging platform into a major source for

* This work is supported by the Natural Science Foundation of Hebei Province of China (Grant No.: F2011203219).
† Corresponding author: Jing Feng (E-mail: kefeng510@163.com).
spreading information of hot events on the Internet\textsuperscript{[2]}. A message of just 140 words sent through microblogs can be the center of public opinions because of the attractiveness of the information itself as well as the dissemination of information in the network.

Users and user relationships, as the nodes and edges of the network, lay the foundation of information transmission and user activities, such as forwarding information pieces, make information transmission possible. Some microblog posts, when initially published, draw little attention, but they can become the center of public attention after being forwarded by some influential users. These influential users play a crucial role in information dissemination. Over the past years, researchers have shown an increasing interest in the study of microblog users’ influence\textsuperscript{[3]}. It is observed that the influential users play a key role in the continuous transmission of the information and the expansion of the transmission scale\textsuperscript{[4]}. These hub users are only potentially influential users. However, if they are not active, it is hard to measure their actual influence.

Sina Weibo, a Twitter-like online service, is one of the most notable and widely used microblogging services in China. Like Twitter, Sina Weibo employs a social networking model called “following”, in which a user is allowed to choose any other users he/she wants to follow without asking any permission. This user is called a “follower” of another user he/she wants to follow. Conversely, the user may also be followed by others without granting permission first. Sina Weibo currently uses the number of a user’s followers to reflect how much attention that this user has drawn from his/her audience. Nevertheless, this method can have problems due to the existence of dummy follower accounts. These dummy followers seldom publish tweets or reply to tweets and remain entirely dormant most of the time.

Based on the study of existing methods of measuring user influence in microblogs, we propose an index model to measure user influence by taking Sina Weibo users as the research sample. We define the audience’s attention that a certain microblogger has drawn from the perspective of link analysis. When analyzing the audience’s attention, we also judge whether or not this microblogger is at the central node in order to measure the potential influence of this microblogger. Combining the audience’s attention that the microblogger has drawn, we will be able to identify who is the influential microblog user, with the level of his/her activity in the network, and that of his/her tweets’ influence.

2 Related work

The initial research on microblog users’ influence was focused on link analysis\textsuperscript{[5]}. Weng et al.\textsuperscript{[5]} proposed TwitterRank, an extension of PageRank algorithm, to measure
the influence of users in Twitter. This approach has taken both the topic similarity between users and the link structure into account. Their algorithm was simple and practicable, but it has problems in objectivity and accuracy.

Cha et al.[6] studied user influence from the aspect of user activities. They proposed three influence measures: Indegree, retweets, and mentions. “Indegree” is the number of people who follow a user, “retweets” mean the number of times the user’s tweet is forwarded, and “mentions” mean the number of times that this user’s name has been mentioned. Link analysis reflects more relations among users, but less the user’s influence. They claimed that it is more influential to have an active audience who retweets or mentions the user. Nevertheless, their influence measures are not sufficient as not only retweets and mentions but also other factors such as the number of responding commentaries reflect the user’s influence.

In order to solve the deficiency of the current user influence ranking mechanism, Shi et al.[7] studied interactions among users and put forward the user activity model to reflect the microblog users’ level of activity. However, their model can only reflect the user’s level of activity, but it cannot fully reflect the user’s influence. Ye and Wu[8] compared and examined the metrics for social influence including the number of followers of a user, replies and retweets, the number of tweets, tweets’ replies and the number of retweets, they pointed out that the number of replies is the most stable metric of the user influence. Li et al.[3] summarized four kinds of methods for evaluating user influence, and improved the TURank algorithm of Yamaguchi et al.[9] according to the characteristics of Chinese microblog systems, but they did not conduct an empirical test of their method, and the effect of their method still remains unknown.

By taking Sina Weibo users as the research sample, we put forward the microblog user influence index model with an emphasis on link analysis and user activities in the microblogging network. To decrease the impact of dummy followers, we used the audience’s attention that a user has drawn to replace the number of followers of the user as a measure of this user’s potential influences. We then calculated the user’s level of activity and his/her tweets’ influence. Based on the two metrics mentioned above, we measure the microblog user’s influence.

3 The user influence index model

A microblog user’s influence is defined as the user’s ability to spread information and to influence others in microblogging systems. It is mainly correlated to the degree of attention that the user can grasp from his/her audience, his/her activities and tweets’ influence. Audience’s attention can be seen as the user’s potential influence from the perspective of link relations, which can effectively reduce the
influence of dummy followers. However, it cannot get rid of the impact of dummy
followers on user influence. Therefore, we combine the impact of audience’s
attention with user activities and tweets’ influence to evaluate the user’s influence.
As the driving force of user influence, user activities reflect the user’s state of how
active he/she is. They are not only correlated to the audience’s attention, but the
frequency of the audience’s increased attention and that of the tweets publication.
The tweets’ influence can reflect the spread scope and the impact of the tweets. It
is correlated to the attention that the user has drawn from his/her audience, the
numbers of retweets and responding commentaries.

3.1 The audience’s attention

Currently, Sina Weibo measures its user influence by the number of the followers
the user has. However, to estimate the user influence by using the size of a user’s
follower count can be misleading as some followers are dummy followers and the
“following” relationship itself does not carry strong indication of influence.
Therefore, our concept of audience’s attention is based on the link analysis, and the
improved PageRank algorithm\[10–12\] we used is to reflect the audience’s attention,
trying to demonstrate a microblog user’s potential influence in a more reasonable
way.

In a microblog system, each user has an independent homepage. A user may have
followers who create links to their accounts in order to receive the content the user
generates automatically. Each follower of the user represents an inbound link,
indicated the support for the user. In the meanwhile, the user may follow others to
receive the content others generate and this “following” may be counted as an
outbound link on the user’s homepage. Analyzing these links on the homepage
reveals users’ relationships in the microblog network.

PageRank is a method for measuring the relative importance of Web pages. When
one page links to another page, it is effectively casting a vote for the other page\[13\].
The basic idea of the improved PageRank algorithm is stated as follows:

An important page is one that is cited by other pages or is capable of being cited
by other important pages. At the same time, the times that the page is cited by others
are much more than those it cites others. A page “votes” an amount of PageRank
onto each page that it links to. Based on the ratio of inbound links against the
outbound links, we can calculate the PageRank (PR) value that this page will receive
from other pages. According to the ratio of the outbound links of this page against
all of its inbound links, we can get different PR values that the page “votes” to other
pages it links to. Our algorithm distinguishes the importance degree of different
pages as inbound links.
A page’s PageRank is calculated as:

\[ PR(u) = (1 - d) + d \sum_{v_i \in P(u)} m_{v_i} PR(v_i) \]

where \( d \) is a damping factor, and usually set to be 0.85. Let \( u \) be a Web page. \( P(u) \) is the set of pages \( v_i \) which points to \( u \), and \( m_{v_i} = \frac{IO_i}{\sum_{i=0}^{n} IO_i} \) is the weight of PR value that \( u \) receives from \( v_i \) linking to \( u \). \( IO_i = \frac{InL_i}{OutL_i} \) defines the amount of PR value that \( v_i \) can receive from the pages that link to it. The greater the value of \( IO \), the easier for \( v_i \) to receive a greater PR value. As it is likely that a Web page does not have any outbound links, its \( IO \) value will be infinite. In this case, we assign a constant \( m_0 \) to the \( IO \) value of a page which has no outbound links. Taking the Chinese famous economist Lang Xianping as an example, his homepage has millions of inbound links but no outbound links. \( InL_i \) is the number of \( v_i \)'s inbound links and \( OutL_i \) is the number of \( v_i \)'s outbound links, representing the amount of PR value that \( v_i \) receives from other pages and \( v_i \) votes onto other pages, respectively.

The weighted calculation of the audience’s attention and iterative process of the matrix will increase the weights of pages with high PR value and reduce the weights of pages with low PR value. Although a dummy user’s page can link to a certain microblogger’s page and confer some importance to this page, as users normally will not pay attention to meaningless dummy users, a dummy user’s page is expected to confer little importance to the page it links to. Our experimental results show that the audience’s attention can not only weaken the impact of dummy users, but also adjust the impact into a proper range, reflecting the actual attention that a user has received from his/her audience (Fig. 1).

Fig. 1 Comparison of the index value of a user’s followers and the audience’s attention.
3.2 The model

User influence indicates the user’s ability to influence others in microblogging systems. It is mainly correlated to the attention that the user has drawn from his/her audience, the user activities and his/her tweets’ influence. In our model, we use the audience’s attention to evaluate a user’s potential influence and the user’s activities and his/her tweets influence to assess the user’s influence.

3.2.1 User activities

A user’s such activities as publishing tweets and following other users to receive tweets can reflect the state of how active he/she is in microblogging systems. A user’s activities will undoubtedly help expand his or her influence in microblogs. The increasing attention that a user received also indicates the expansion of his/her influence. Different from Shi et al’s study, instead of using the number of a user’s followers, and the frequencies with which a user follows others and the user publishes tweets to build a user activity index model[7], we used the audience’s attention to replace the number of the user’s followers, and regarded a user’s following others and publishing tweets as the influence measures of the user’s level of activity. We assign different weights to the two measures and the level of activity is calculated as:

$$UV_i = \sum_{j=1}^{k} \gamma_{ij} x_{ij} + UA_i$$  \hspace{1cm} (2)$$

where $UV_i$ is the level of activity of the user $i$, $UA_i$ is the level of the audience’s attention that the user $i$ has drawn, $x_{ij}$ is the $j$th influence factor of the level of the user $i$’s activities, and $k$ is the number of influence factors and $\gamma_{ij}$ is the weight of the $j$th influence factor of the user $i$.

$$x_{ij} = \frac{X_{ij}}{T_{first,j} - T_{last,j}}$$  \hspace{1cm} (3)$$

where $T_{first,j}$ is the latest occurring time of the $j$th impact factor of the user $i$, which means the latest time for this user to post messages or to follow a concerned user. $T_{last,j}$ is the earliest time to post messages or to follow a concerned user. $T_{last,j} - T_{first,j}$ indicates the total days of one impact factor once it occurred. $X_{ij}$ is the total number of the $j$th impact factors, and $x_{ij}$ is the frequency with which the $j$th influence factor of the user $i$ occurs.

3.2.2 Tweets’ influence

We measure the influence of a user’s tweets mainly by the number of retweets and the number of responding commentaries. The more retweets and commentaries that a user receives, the more attention the audience pays to the tweets and the greater
influence of the tweets. The tweets’ influence can be quantified by the audience’s attention, the numbers of retweets and responding commentaries. Based on different contributions of the three metrics, we calculate a tweet’s influence as:

$$MI_{i,j} = UA_i + 3\sqrt{MR_{t,i,j}} + \sqrt{MC_{i,j}}$$  (4)

where $MI_{i,j}$ is the $j$th tweet’s influence of the user $i$, $UA_i$ is the attention that user $i$ has received from the audience, $MR_{t,i,j}$ and $MC_{i,j}$ are the numbers of retweets and commentaries on the $j$th tweet of the user $i$, respectively.

The influence of the user $i$’s tweets is defined as the average of the influence of all tweets published by the user $i$. It is calculated as:

$$MI_i = \frac{\sum_{j=0}^{n} MI_{i,j}}{n}$$  (5)

where $MI_i$ is the tweets’ influence of the user $i$ and $n$ is the total tweets that the user published.

3.2.3 User influence

The user’s influence is measured by combining the user’s level of activity with tweets’ influence, which is calculated as:

$$UI = \alpha \cdot UV + (1 - \alpha)MI$$  (6)

where $\alpha$ is the balance factor to adjust the weight of $UV$ and $MI$.

4 The experiment and the result analysis

4.1 Data collection

To investigate the performance of our model, we implement experiments using data crawled from Sina Weibo. Data and the related data relations are summarized as follows:

- Attributes of users. User ID, user categories, degree of the audience’s attention, the number of followers, and the number of tweets.
- Attributes of tweets. The number of tweets, the time to publish tweets, the number of retweets, and the number of commentaries.
- Relationship of users. User ID, ID of other users that this user follows, and ID of followers of this user.

The data collection work began in November of 2011 and ended on Mar. 5 of 2012. According to the requirement of the model, we collected data between Feb. 5 and Mar. 5 of 2012 on 5 microblog users.
4.2 The result analysis

Different weights of the influence factors are calculated through analytic process hierarchy (APH), where in computing the user’s level of activity, \( \gamma = \{0.2025, 0.1778, 0.6197\} \), and in computing user influence, \( \alpha = 0.397 \). The values of different influence measures of 5 users are listed in Table 1.

<table>
<thead>
<tr>
<th>User</th>
<th>Attention received</th>
<th>User activity</th>
<th>Tweets’ influence</th>
<th>User influence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.1221</td>
<td>2.6749</td>
<td>80.80</td>
<td>49.55</td>
</tr>
<tr>
<td>2</td>
<td>2.8740</td>
<td>5.4124</td>
<td>11.66</td>
<td>9.16</td>
</tr>
<tr>
<td>3</td>
<td>6.6162</td>
<td>4.8000</td>
<td>55.37</td>
<td>35.14</td>
</tr>
<tr>
<td>4</td>
<td>3.5740</td>
<td>0.8913</td>
<td>64.13</td>
<td>38.83</td>
</tr>
<tr>
<td>5</td>
<td>5.4330</td>
<td>2.4958</td>
<td>95.15</td>
<td>58.09</td>
</tr>
</tbody>
</table>

The experimental results are illustrated in Fig. 2. Figure 2(a) shows the relationship between the audience’s attention, a user’s potential influence and the user’s actual influence. Generally speaking, changes of the user’s potential influence were...

Fig. 2 Relationship between audience’s attention and user influence (a). The impact of user activities and tweets’ influence with equal (b) and different (c) weights on user influence.
basically consistent with those of the user’s influence. However, an abnormal change was observed about the fourth user. Normally, the more potential influence a user has, the more influential the user will be. But Fig. 2(a) shows that the fourth user with high potential influence has a low actual influence. This indicates that the audience’s attention shows the user’s potential influence rather than the user’s actual influence. This result is in line with the conclusion of Cha et al’s study[6].

Figure 2(b) shows the impact of user activities and tweets’ influence with equal weight on user influence. Obviously, tweets’ influence produces a greater impact than user activities on the user’s actual influence. This result also proved that microblog users increase their influence through posting tweets. We assigned different weights to user activities and tweets’ influence. Figure 2(c) demonstrates the result of the impact of the two measures with different weights on the user’s actual influence.

5 Conclusion

The paper studied the user influence ranking mechanism in microblogs. We measured user influence from the perspective of link analysis and user activities. Firstly, we proposed the concept of the audience’s attention and used it to replace the number of followers of the user to indicate the user’s potential influence. We used the link analysis to measure the audience’s attention, which reflects the attention that a user can garner from his/her audience and the user’s potential influence in a more reasonable way. Then, we defined the user level of activity and the tweets’ influence to assess the user’s influence. The experimental results show that our user influence model performs well and it can reflect the users’ real influence in a more reasonable way.

User influence evolves as time goes by. Our model should be a dynamic one, which can distinguish the changes of user influence at different time periods. So in our future research, we will add more influence measures and assign the proper weights of each measure to identify different influential users at different time periods.

References

Contents Index to Volume 5

Perspective
Transforming the academic library for the future

Research Papers
An empirical research on strategic planning in public libraries of Mainland China

Information behaviors of Chinese K-12 physical education teachers:
A survey study

Kewen WU, Qinghua ZHU, Julita VASSILEVA & Yuxiang ZHAO 2012,5(1):34–50

Investigating the relationships between facets of work task and selection and query-related behavior

A preliminary study on exploratory search behavior of undergraduate students in China
Yunqiu ZHANG, Wenxiu AN & Jia FENG 2012,5(1):70–84

Defining an open access resource strategy for research libraries:
Part II—Challenges and responses

Academic institutional repositories in China: A survey of CALIS member libraries
Yongchao ZHAO, Xiaoxia YAO & Chengfu WEI 2012,5(2):18–32

Cross-efficiency DEA model-based evaluation of allocative efficiency of rural information resources in Jiangsu Province, China

Integrating innovative library e-reserves service into teaching and learning at Shanghai Jiao Tong University
Xiaohua SHI, Xiaowei YU & Wei PAN 2012,5(2):45–58

Information literacy course design based on student survey:
The practice of subject librarians at NSL, CAS
Ming WU, Li WANG & Yanli LIU 2012,5(2):59–73
A study of the effects of students’ cognitive styles on their learning behaviors of online database search strategies
Ying FENG, Liren GAN & Deyi QIAO 2012,5(2):74–86

Defining an open access resource strategy for research libraries: Part III—The strategies and practices of National Science Library
Xiaolin ZHANG, Xiwen LIU, Lin LI, Yan ZENG & Li-Ping KU 2012,5(3):1–11

A view on big data and its relation to Informetrics
Ronald ROUSSEAU 2012,5(3):12–26

Person-specific named entity recognition using SVM with rich feature sets
Hui NIE 2012,5(3):27–46

A comparison of mapping strategies from DDC to CLC

A study of information exchange through social networks in rural China
Ya LIU 2012,5(3):62–75

Is it time for wider acceptance of e-textbooks? An examination of student reactions to e-textbooks
Ziming LIU 2012,5(3):76–87

A knowledge discovery method based on analysis of multiple co-occurrence relationships in collections of journal papers

Designing an automated FAQ answering system for farmers based on hybrid strategies

Factors influencing knowledge contribution: An empirical investigation of social networking website users
Shengli DENG, Ting ZHOU & Min ZHANG 2012,5(4):37–50

A comparative study of information diffusion in weblogs and microblogs based on social network analysis

An index model for measuring microblog users’ influence
Fuyong YUAN, Jing FENG & Qianqian FU 2012,5(4):67–76

A method for improving the accuracy of automatic indexing of Chinese-English mixed documents
Yan ZHAO & Hui SHI 2012,5(4):77–92

Research Progress
Digital reading: An overview
Chinese Journal of Library and Information Science

1 基本信息

期刊名称：Chinese Journal of Library and Information Science (CJLIS)
创办时间：2008年
主管单位：中国科学院
主办单位：中国科学院国家科学图书馆（即中国科学院文献情报中心）
出版单位：《中国文献情报》（英文刊）编辑部
主编：张晓林教授（中国科学院国家科学图书馆馆长）
刊号：ISSN 1674-3393；CN 115670/G2
每期定价：国内：200元/期（800元/年、卷）；国际：199美元/年、卷（含邮资）

2 期刊定位与报道范围

国内图书情报领域唯一一份面向国内外公开征订与发行的学术性英文期刊，是国内图书、情报与档案学领域的国际化高端学术交流平台与工具。刊载内容涵盖图书馆学、情报学、档案学与新在图书馆的中的应用，包括图书馆学情报学教育等相关研究领域的各种理论研究与技术实践。

3 论文类型与评审

CJLIS注重论文的前沿性、学术性与实用性。秉持理论与实践并重，采用同行匿名评议。其编委会由海内外23位知名专家与学者组成，其中，境外编委约占2/3。CJLIS主要发表原创性研究，重点为实证性研究论文，包括研究综述与实践研究等学术论文。

4 读者对象

CJLIS是国内外图书情报机构从业者、广大图情研究管理人员、教师、硕博研究生和本科生的重要参考工具，是了解国内外文献情报事业发展现状和最新研究成果的重要窗口。于每季末25日出版。年出版4期。如欲订阅，可通过以下方式办理：

1. 邮局订阅：全国各地邮局均可办理
   - 刊号：ISSN 1674-3393；CN 11-5670/G2
   - 邮发代号：82-563

2. 国际代理
   - Ebsco International, Inc. (USA)

3. 《中国文献情报（英）》编辑部
   - 地址：北京海淀区中关村北四环西路33号
   - 5067房间
   - 邮编：100190
   - 电话：010-826224456；010-82626611转6628
   - 传真：010-82622445
   - 联系人：曹菁
   - E-mail：chinalibraries@mail.las.ac.cn

4. 北京中科期刊出版有限公司
   - 地址：北京西城区德胜门外大街83号
   - 德胜国际中心B座3层301室
   - 邮编：100088
   - 电话：010-59790736；8028
   - 传真：010-59790736-8092
   - 联系人：孙志蓬
   - E-mail：sunzhipeng@mail.sciencep.com

5. 中科院国家科学图书馆编辑出版中心发行组
   - 地址：北京海淀区中关村北四环西路33号
   - 邮编：100190
   - 电话：010-82622346
   - 传真：010-82626611转6155
   - 联系人：苗志刚
   - E-mail：miaozg@mail.las.ac.cn
Chinese Journal of Library and Information Science

(CJLIS, Chinese Title: 《中国文献情报》)

ISSN 1674-3393, CN 11570/G2; Price: $US 199/Volume

Chinese Journal of Library and Information Science (CJLIS) is the first and only English-language academic journal in the field of Library and Information Science (LIS) that has been published in Mainland China so far. It aims to establish a platform for students, researchers and professionals in LIS and archive science to engage in intellectual dialogs and to share best professional practices of their experience.

CJLIS adopts a double-blind peer-review process for each published article. The Editorial Board of CJLIS is composed of well-known scholars both from national and international LIS fields. The main aim of CJLIS is to serve librarians, information professionals, researchers and practitioners in the LIS field for their various purposes and keep them well informed of the latest development of LIS field both in China and the world.

CJLIS mainly publishes original research articles, including research papers, library practice papers, progress reports, research reviews and perspectives. For subscription, you can use the following channels:

1) Via EBSCO International, Inc (USA)

2) To order at any post office inside China, with a distribution code 82-563.

3) To contact with the Editorial Office
   Address: No.33 Beisihuan Xilu, Zhongguancun, Haidian District, Beijing 100190, P.R. China
   Tel: 86-10-82624454 or 86-10-82626611 ext. 6628
   Fax: 86-10-82624454
   Email: chinalibraries@mail.las.ac.cn
   Website: http://www.chinalibraries.net

4) Via Beijing China Science Journal Publishing Co., Ltd.
   Address: Room 301, Building B, Desheng International Center, No. 83 Deshengmenwai Street, Beijing 100088, P.R. China
   Tel: 86-10-59790736 ext. 8028
   Fax: 86-10-59790736 ext. 8092
   E-mail: sunzhipeng@mail.sciencep.com
Submission Guidelines

◆ Aims
Chinese journal of Library and Information Science (CJLIS), being sponsored by the Chinese Academy of Sciences (CAS) and published quarterly by the National Science Library of CAS, is a scholarly journal in the field of library and information science (LIS). Its aim is to provide an international communication link between researchers, educators, administrators, and information professionals.

With the publication of the research results both from China and from other foreign countries, the Journal CJLIS strikes a balance between theory and practice. With its goal to provide an open forum for Chinese and international scholars in this field to exchange their research results, CJLIS also offers new possibilities in the advancement of Chinese library operations. The CJLIS tries to establish a platform for LIS students, researchers and library staff all over the world to engage in intellectual dialog and also to improve library services so as to promote even more quickened and substantial development of LIS in China.

◆ Scope
Striving toward academic excellence, innovation, and practicability, the CJLIS mainly includes research papers both on the theoretical as well as on the practical fronts in all aspects of the field. More specifically, it includes but not limited to informatics, library management, information technology application, knowledge organization system, knowledge management, archives, permanent preservation of library resources, LIS education, and so on.

◆ Refereeing Process
Articles and papers covering the topics or themes mentioned above will be refereed through a double-blind peer review process.

◆ Editorial Advisory Board
The Editorial Board is composed of the nationally and internationally well-known scholars and researchers in the LIS field and the high quality of this Journal is thus reasonably assured.

◆ Manuscripts Categories
As the first English-language academic journal on LIS published in Mainland China, the CJLIS will take a proactive attitude to trace and report the prevailing hot issues in the field around the globe as well as the more serious scholarly communications. As such, the submitted manuscripts are classified into constant categories and unfixated categories. In the former category, research papers, library practice and progress reports are the essential components. In the latter, book reviews, biographical sketches, anecdotes, reminiscence of prominent librarians and brief communications will appear occasionally.

Research papers represent original research work or a comprehensive and in-depth analysis of a topic. More than 3,000 words are considered as a proper length for such manuscripts, with a structured abstract ca. 200 words.

Library practice covers the latest development and application in any segment of library field work and information service. The length of the manuscript is preferred to be more than 3,000 words, with a structured abstract ca. 200 words.

Progress reports reflect the projects result or research progress on the key topics of the library and information science. Submissions of articles to this section are expected to be comprehensive and analytical, which may deepen the understanding of the discussed issue and stimulate further researches on the topics, or give a new perspective on future technological applications. The manuscript length should be within 5,000 words, with a structured abstract ca. 200 words.

◆ Manuscripts Requirements
All papers can be submitted either in English or in Chinese (or both) with a double line space. For the assurance that all the materials of the to-be-submitted are included, please check the following:

Title. Please give a brief biographical introduction to all contributing authors and their research background on a separated paper. For a better organization of the paper, please use the headings and subheadings.

Authors and affiliations. Please do not forget to write down the mailing address of each and every article contributors.

References. Be sure all the references used should be cited properly in both in-text and in bibliography. Particular attention should be paid to the proceedings. Do not forget adding the name(s) of editors of the compilation, as well as the name of the publishers. For the detailed information, please request a copy of Reference Citation Format.

◆ Copyright
All submitted papers normally should not have been previously published nor be currently under consideration for publication elsewhere. For all the materials translated or obtained from other published resources, they should be properly acknowledged. All copyright problems should be cleared without any legal entanglements prior to the publication.

◆ Notes for Intending Submissions
A guide for authors and other relevant information, including submitting papers online, is available at the website of the Editorial Office of the CJLIS (http://www.chinalibraries.net). For any questions, you can e-mail the Office or directly to:

Prof. ZHANG Xiaolin
Editor-in-Chief of CJLIS
The CJLIS Editorial Office
National Science Library, Chinese Academy of Sciences
No.33 Beishihu Xiliu, Zhongguancun, Haidian District,
Beijing 100190, P.R. China
Tel: 86-10-82624454 or 86-10-82626611 ext. 6628
Fax: 86-10-82624454
E-mail: chinalibraries@mail.las.ac.cn
Website: http://www.chinalibraries.net

◆ Subscription
For single-copy subscription in China, RMB ¥ 200/Issue. For subscription outside of China, US $ 199/Volume yearly (including air shipping)
Contents

Perspective
1 Transforming the academic library for the future
   Gary E. Strong

Research Papers
9 A knowledge discovery method based on analysis of multiple co-occurrence relationships in collections of journal papers
   Hongshen Pang

21 Designing an automated FAQ answering system for farmers based on hybrid strategies
   Juntian Zhang, Xuefang Zhu
   & Guang Zhu

37 Factors influencing knowledge contribution: An empirical investigation of social networking website users
   Shengli Deng, Ting Zhou & Min Zhang

51 A comparative study of information diffusion in weblogs and microblogs based on social network analysis
   Yang Zhang & Wanyang Ling

67 An index model for measuring microblog users’ influence
   Fuyong Yuan, Jing Feng & Qianqian Fu

77 A method for improving the accuracy of automatic indexing of Chinese-English mixed documents
   Yan Zhao & Hui Shi

93 Reference Citation Format

Content Indexes
94 Contents Index of Volumes 1–4
101 Contents Index of Volume 5

Enlightment to Subscribe
103 Print Subscription