The Analysis of Online News Information Credibility Assessment on Weibo Based on Analyzing Content

Quan Yuan(✉) and Qin Gao

Department of Industrial Engineering,
Tsinghua University, Beijing 100084, China
yuan-ql5@mails.tsinghua.edu.cn,
gaoqin@tsinghua.edu.cn

Abstract. As a representative of online news information carrier, social media contain much information that difficult to be differentiate between true and false. Sina Weibo is the most popular social media in China and need to be took into account. This paper first identify criterions and cues for assessing the credibility of online news through literature review. Second, we interviewed 5 experts in the news information judgment process. According to the interview, we reviseed the effective criteria and cues for helping users assess the online news information. Third, we design and develop an assistant webpage as a tool to help Weibo users assess the credibility a specific Weibo news by visualizing cues related to its source, content, dissemination, and topic. Future work are to test the effectiveness and efficiency of the tool and improve it.

Keywords: Information credibility · Online news · Sina Weibo · Assistant assessment · Content analyzing

1 Introduction

As technology rapidly develops, the Internet has already enter into worldwide families of millions. The convenient Internet not only brings people information without leaving home, but also makes people easier to send messages to other people: People can easily communicate with others immediately without the constraint of distance through online chart tools, can read current news all over the world at first time when the event occur through web news portals, and can speak their words out letting others see them through social network.

However, as everyone can be the publisher of the information, Internet also brings many rumors and bogus news to people, which look so real that most of us cannot tell the truth. As online information circulate quickly, once a rumor tend to spreads in many people, it is very likely to cause huge losses for people. Thus, we should pay attention to the information credibility problems in the Internet.

Like Facebook and Twitter, Sina Weibo is one of the representative new social media in China. Weibo has the significant influence on Chinese. In June 2015, there exists 212 million monthly active users in Weibo. So it is absolutely an important social platform

© Springer International Publishing Switzerland 2016
D. Harris (Ed.): EPCE 2016, LNAI 9736, pp. 125–135, 2016.
DOI: 10.1007/978-3-319-40030-3_14

which can be the source of Internet rumors. In March 3, 2011, Fukushima Nuclear Power Plant leaked because of the big earthquake. In March 16, a fake news claiming the leak had affected China, and iodate salts could prevent the damage spreads on the Internet, particularly in Weibo, in an explosive speed. Most of Internet users believed it, buying iodate salts crazily. The stock around China was insufficient, and it seriously impacted the lives of Chinese. Once a rumor is read by users, it can mislead the truth to another direct, causing damage in many aspects. It is necessary for us to search for a way to assist Internet users in judging online information credibility, particularly in Weibo.

The objective of our research is to propose a method which can guide Internet users to the scientific approach to judging the credibility of information in the Internet according to the content of information, and develop an assistant webpage with the method based on Weibo, and conduct an experiment to verify the effectiveness of the method and the usability of the website.

2 Judgment Process, Criterions and Cues in Judgment

Information credibility is what degree a person believe the information is true when receiving it. It is content-based, but different from different people and different environment, and easily affected by kinds of reasons.

2.1 Judgment Process

Before people judging the credibility of online information, they generally make sure the topic of the information is interesting for them [1]. They begin the judging process in judge whether the website providing the information is authoritative and credible [2]. People continue the credibility judging process only if they trust on the platform. Rather than accuracy, people usually expect more efficiency when judging credibility [3], and people have limited capacity in mind [4]. Thus people will not spend much time on it, but prefer to use a small part of information they have, such as words, pictures and other cues, [5] for judging in terms of some heuristic including reputation, endorsement, consistency, expectancy violation and persuasive heuristic [6] and one or two criterions [7] from usefulness, completeness, accuracy, currency, and importance. [8] Consequently, people sometimes cannot do a correct judgment to an information.

If the motivation for judging is not enough for getting conclusion, people will give up. Even if people have their own conclusion, because of the shortage of process, there may be the misjudgment about judging information credibility.

2.2 Judging Criterions

Metzger reviewed current recommendations helping people judge credibility of information, and evaluated the existing cognitive models for credibility assessment, and provided the future online credibility education and practice [7]. He suggested five criterions for judging process including accuracy, authority, objectivity, currency and coverage. However, Internet users always follow accuracy, authority, objectivity and coverage, but ignore currency. When the time and energy for assessment are limited, they will just use two in five criterions, e.g., accuracy and coverage, and even use only one criterion for the final decision. When user judge the credibility, they are always partial to beautiful interface. For assisting users, we can emphasize criterions that are easier to be ignored for them.

Table 1 is the criterions with its details and definitions.

Table 1. Judging criterions in the process of information credibility assessment [7, 9]

Judging criterions	Short definitions
Accuracy	Information provides a true description of reality
	Information is trustable, giving the same result on successive trials
	Information is able to accurately describe reality
Authority	Information is presented in authorized forums
	Information is based on the findings of scientific research
Objectivity	A piece of information is presented as an objective description of reality
	Information provides an impartial and unbiased description of reality
Currency	Information is timely, recent, or up-to-date
	Information provides multi-aspect description of current reality
	Information provides something really new
Coverage	Information covers a broad range of facts and opinions
	Information is focused enough to match the needs of a person or a group
	Information is considered as helpful to meet the need of a person or a group

2.3 Judging Cues

Schwarz and Morris suggested that providing for Internet users some cues that can supporting to evaluate the credibility of online text was useful and valuable [10]. Castillo, Mendoza and Problete classified cues that influence people's judgment of online information credibility into information contents, information sources, topics, and information dissemination [11]. In addition, Flanagin and Metzger considered that visual styles of information also intensively affected the assessing of credibility, [12] and Fogg even believed that the influence of visual styles was more important than information content and source [3].

Information Contents: Properties related only to the content of information, including misspelling, misuse of grammar and punctuation, reference of literature or data, professional and clear content, understandability, reasonability, advertisement, URL, tags, and so on [7, 10, 13].

Information Sources: Properties about the publishing sites and author of online information. For sites, there are many cues that can affect the credibility of information it publishing, including visitor volume, wining prizes, ranking, official, domain type, and simple navigation, structural websites, notification of review from editor, interaction function (search, information confirming, prompt response for user service), reputation [7]. The cues about author influencing the judgment are gender, geographic location, active degree, expert in topic area, officially proving, and contact details [13, 14]. Specially, in social platform, the head portrait, followers and followees are also important for users assessing the credibility [13, 15].

Topics: The topic about the information content and its properties. Currencyis the most obvious cues of topic [7].

Information Dissemination: Description of how the information dissemination. Popularity in experts, visitor volume classified by geographic location, in the front of the list of search results, sharing times are all the properties [10, 13].

Visual Styles: Properties of the ways the information visually present to users, independent with other properties of information itself. Personalizing homepage, [13] professional, attractive, consistent design of sites is included in visual styles.

Table 2 shows affecting cues about online information credibility.

3 Expert Interviewing

Considering the importance of motivation and people are usually interested in news, in order to make sure which criterions and cues have positive effects on judging outcome, and to add some criterions and cues that may not be mentioned in literatures, interviews with expert in judging credibility of news information were conducted.

3.1 Interview Process

Participants: Five experts in news information credibility assessment, including four Chinese editors of newspapers and one Journalism professor from Tongji University, three male and two female. One editor works for academic journal of Beijing Institute of Technology, and two works for daily news now, one has worked for daily news. They are occupied in jobs related to news, and are all experienced in the credibility about news information.

Interview Questions: Including some open end explorative questions about experts' experience in news credibility assessment, and a questionnaire inquiring the experts' ranking for the positive effect on five criterions and five cues mentioned above. Main aspects of questions were experts' opinions about online news information credibility, the judging process expert usually use in their work, the properties of fake news, and how to help general people assess online news credibility. And in the questionnaire, the 7-point Likert scales were used to measure different levels of agreement to the positive effect of items including criterions and cues from "1 = totally disagree" to "7 = totally agree". The definition of items were provided in the questionnaire.

Procedure: One interview was conducted face to face, and other four interviews was conducted by online video chatting. The interview time for each participant was

Table 2. Cues in the process of information credibility assessment

Type of cues	Cue	Definite items
Information contents	Language errors	Misuse of grammar and punctuation
		Misspelling
	Special contents	URL
		Referrences
		Advertisement
	Quality of content	Professional and clear content
		Comprehensibility
		Reasonability
Information sources	Popularity of the website	Visitor volume
		Prizes
		Ranking
	Professionality of the website	Official platform
		Reputation
	Design and function of the website platform	Simple navigations and structural website
		Notification of review from editors
		Interaction function
	Personal information of authors	Gender
		Contact information
	Properties of authors	Active degree
		Expert in information topic area
		Officially proved
Topics	Currency	Lag between topic event occurring and information published
		Lag between topic event occurring and users seeing it
	Evaluations to the topic event	Opinions in the original website platform
		Opinions in other online platforms
Information dissemination	Event topic spread	Popularity in experts
		In the front of the list of search results, Sharing times
Visual styles	Visual design	Personalizing homepage
		Professional design
		Consistency of the website

40 ~ 90 min. We first asked open questions to experts about their works and specific cases for judging news credibility they had experienced, then let them fill in the questionnaire about useful criterions and cues.

3.2 Interview Results

Properties of Online News: Compared to news on traditional media, online news cares more about ‘new’, and is required to be faster by people. At the same time, the examine steps of online news is simpler and fewer than traditional news. Even so, official news is still more credible than we media on the Internet. Judging credibility of online news only in terms of content is extremely difficult for everyone.

Types of Fake News: There are three different fake news. The first one is completely fake news, representing the news whose content is absolutely fake, such as the medicine that guarantee to cure all diseases. The second fake news is partial fake news, representing the news that contain true part and fake part, which are common on the Internet. The third fake news is bias news. The bias news include no fake part, but hide some fact or enlarge some fact in content. It is the type of majority fake news, causing most damage to people.

Suggestions to Internet Users: Do not easily believe online news. Source, experts’ opinion and integrity of news are all important for credible news. The news whose related news are from only one source is not credible.

Revise of Criterions: According to the questionnaire data, items that mean value of point are greater than 4 and variance of point are small are keep. Useful criterions includes accuracy, authority and objectivity.

Revise of Cues: According to the questionnaire data, items that mean value of point are greater than 4 and variance of point are small are keep. Useful type of cues were information contents, information sources, topics and information dissemination.

4 Assistant Webpage Tool Design

An assistant webpage is designed to help people assess online news in Weibo. The webpage is developed by Python, PHP, JavaScript, CSS and HTML, visualizing the useful cues to Weibo users. Figure 1 illustrates a part of webpage.

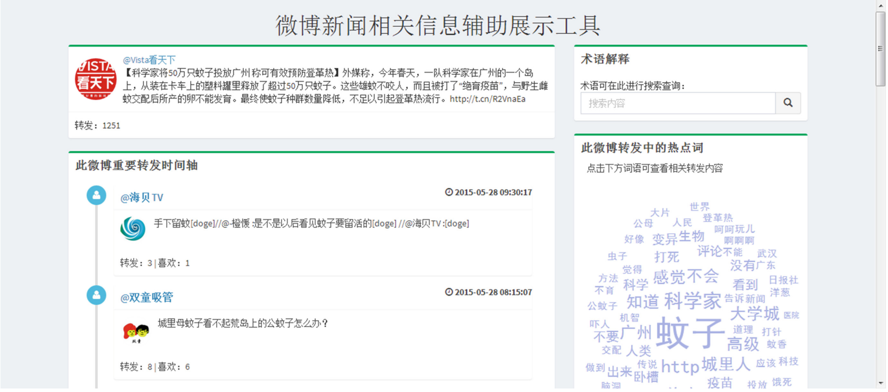
According to the different types of cues, the assistant webpage is divided into three parts, presenting information contents and sources, information spread, and topics. Every part is constituted by different functional pieces, and each functional piece visualizes different cues respectively.

Table 3 shows the structure of webpage with affecting cues.

4.1 First Part: Information Contents and Sources Part

First part contained the information contents and sources assistance. And this part is design as familiar with Weibo as possible considering the consistency for users.

There are two functional pieces in the first part. The left piece presents the content of Weibo news and information of its author, including author’s head portrait, followers, followees, numbers of weibos, active degree and related field, showed by the


微博新闻相关信息辅助展示工具
此微博重要转发时间轴
术语解释
术语可在此进行搜索查询：
搜索内容
此微博转发中的热点词
点击下方词语可查看相关转发内容
世界
大片
登革热
公母
人民
呵呵玩儿
好像
变异
生物
啊啊啊
虫子
打死
评论
不能
武汉
觉得
没有
广东
方法
科学
感觉
不会
看到
日报社
不育
洋葱
公蚊子
知道
科学家
告诉
新闻
大学城
医院
机智
吓人
道理
打针
不要
广州
蚊子
高级
蚊香
交配
人类
做到
传说
http
城里人
应该
科技
出来
卧槽
脑洞
疫苗
投放
饿死

@Vista看天下
【科学家将50万只蚊子投放广州 称可有效预防登革热】外媒称，今年春天，一队科学家在广州的一个岛上，从装在卡车上的塑料罐里释放了超过50万只蚊子。这些雄蚊不咬人，而且被打了"绝育疫苗"，与野生雌蚊交配后所产的卵不能发育。最终使蚊子种群数量降低，不足以引起登革热流行。http://t.cn/R2VnaEa

转发：1251

@海贝TV 2015-05-28 09:30:17
手下留蚊[doge]//@-橙慏 :是不是以后看见蚊子要留活的[doge] //@海贝TV :[doge]
转发：3 | 喜欢：1

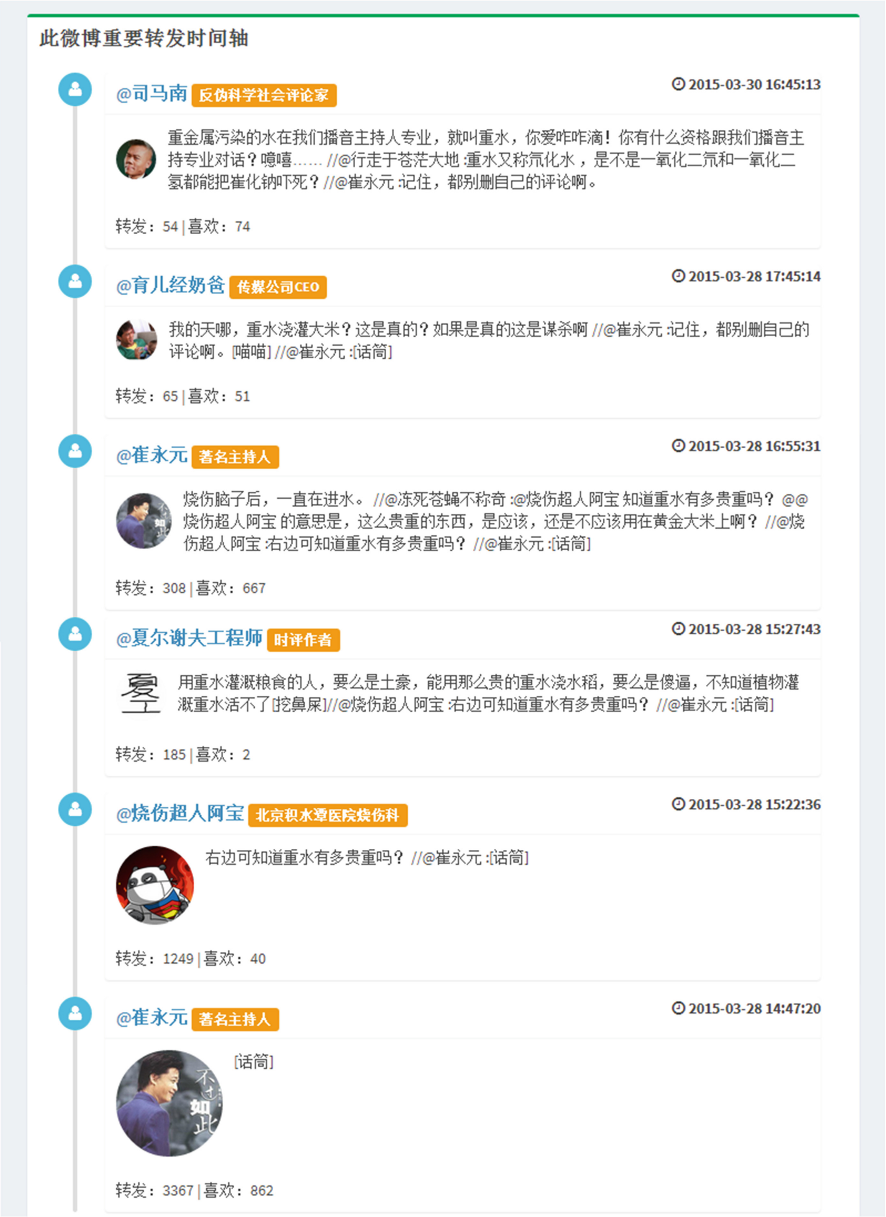
@双童吸管 2015-05-28 08:15:07
城里母蚊子看不起荒岛上的公蚊子怎么办？
转发：8 | 喜欢：6

Fig. 1. The overall design of the assistant webpage

Table 3. The structure and content of the assistant webpage

Type of cues	Cues	Definite items	Webpage part	Functional piece
Information contents	Quality of the content	Professional and clear content	First part	Explanations of terms in the information
		Reliable data		
Information sourses	Properties of authors	Expert in news topic area	First part	Authors' Keywords
		Officially proved identity		Identification in Weibo
Topics	Currency	Following up the development of the information topic	Second part	Retransmission timeline
	Evaluations to the topic event	Opinions in the website	Third part	Relevant news in Weibo
		Opinions in other platforms		Relevant news in other websites
Information dissemination	Event topic spread	Opinions of experts	Second part	Retransmission timeline
				The clouds of hot words

rate of sending weibos and keyword of author's weibos. The right piece presents the means of terms in news content from Wiki, helping users understand the news.



此微博重要转发时间轴

@司马南 反伪科学社会评论家 2015-03-30 16:45:13

重金属污染的水在我们播音主持人专业，就叫重水，你爱咋咋滴！你有什么资格跟我们播音主持专业对话？嘻嘻…… //@行走于苍茫大地 :重水又称氘化水，是不是一氧化二氘和一氧化二氢都能把崔化钠吓死？//@崔永元 :记住，都别删自己的评论啊。

转发：54 | 喜欢：74

@育儿经奶爸 传媒公司CEO 2015-03-28 17:45:14

我的天哪，重水浇灌大米？这是真的？如果是真的这是谋杀啊 //@崔永元 :记住，都别删自己的评论啊。[喵喵] //@崔永元 :[话筒]

转发：65 | 喜欢：51

@崔永元 著名主持人 2015-03-28 16:55:31

烧伤脑子后，一直在进水。 //@冻死苍蝇不称奇 :@烧伤超人阿宝 知道重水有多贵重吗？ @@烧伤超人阿宝 的意思是，这么贵重的东西，是应该，还是不应该用在黄金大米上啊？ //@烧伤超人阿宝 :右边可知道重水有多贵重吗？ //@崔永元 :[话筒]

转发：308 | 喜欢：667

@夏尔谢夫工程师 时评作者 2015-03-28 15:27:43

用重水灌溉粮食的人，要么是土豪，能用那么贵的重水浇水稻，要么是傻逼，不知道植物灌溉重水活不了[挖鼻屎]//@烧伤超人阿宝 :右边可知道重水有多贵重吗？ //@崔永元 :[话筒]

转发：185 | 喜欢：2

@烧伤超人阿宝 北京积水覃医院烧伤科 2015-03-28 15:22:36

右边可知道重水有多贵重吗？ //@崔永元 :[话筒]

转发：1249 | 喜欢：40

@崔永元 著名主持人 2015-03-28 14:47:20

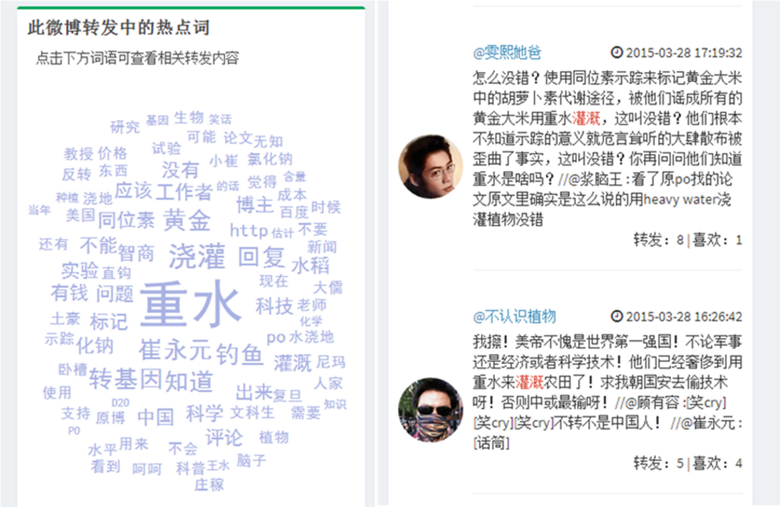
[话筒]

转发：3367 | 喜欢：862

Fig. 2. The timeline functional piece in the second part of the assistant webpage

4.2 Second Part: Information Dissemination Part

Second part of assistant webpage contained the information dissemination of Weibo news. The spread is by ‘retransmit weibos’ function, like retweet in Twitter.


此微博转发中的热点词

点击下方词语可查看相关转发内容

研究 基因 生物 笑话
可能 论文 无知
教授 价格 试验
氯化钠 小崔
反转 东西 没有
含量
应该 工作者 的话 觉得
种植 浇地 成本
当年 博主 时候
美国 百度
同位素 黄金
http 估计 不要
还有 不能 智商 浇灌 回复 新闻
水稻
实验 直钩
现在
有钱 问题 重水 大儒
科技 老师
土豪 标记
化学
示踪 化钠 po 水浇地
崔永元 钓鱼
灌溉 尼玛
卧槽
转基因 知道 人家
使用 出来 复旦
D20 知识
支持 原博 中国 科学 文科生 需要
PO
评论 植物
水平 用来 不会
看到 呵呵 科普 王水 脑子
庄稼

@雯熙她爸 2015-03-28 17:19:32

怎么没错？使用同位素示踪来标记黄金大米中的胡萝卜素代谢途径，被他们谣成所有的黄金大米用重水灌溉，这叫没错？他们根本不知道示踪的意义就危言耸听的大肆散布被歪曲了事实，这叫没错？你再问问他们知道重水是啥吗？//@浆脑王 :看了原po找的论文原文里确实是这么说的用heavy water浇灌植物没错

转发：8 | 喜欢：1

@不认识植物 2015-03-28 16:26:42

我擦！美帝不愧是世界第一强国！不论军事还是经济或者科学技术！他们已经奢侈到用重水来灌溉农田了！求我朝国安去偷技术呀！否则中或最输呀！//@顾有容 :[笑cry][笑cry][笑cry]不转不是中国人！ //@崔永元 :[话筒]

转发：5 | 喜欢：4

Fig. 3. The cloud of key words functional piece in the second part of the assistant webpage

The second part also contains two functional pieces. The left piece is the timeline of key retransmission of the Weibo news, called retransmission timeline. Hot retransmissions from officially proving users and their identity in Weibo are presented ordering by time, gathering the experts' opinion in it. Figure 2 shows the details of retransmission timeline piece.

The right functional piece is the cloud of hot words of all the retransmission by Weibo users. It contains the words that are mostly mentioned in retransmission, and displays the content of related retransmission when clicking a word. The cloud of hot words represent the opinion of general users in spreading process. Figure 3 is the details of the piece.

4.3 Third Part: Topics Part

Third part of assistant webpage shows the cross-platform relevant news about the same topics. It has two functional pieces presenting relevant news in Weibo and at other news websites respectively.

5 Conclusion

In this study, a method of helping the Internet users judge the credibility of online news information was proposed through literature reviewing and interviews with experts in news credibility assessment. The method contained a judgment process, three judging criterions and four affecting cues. The judgment process described how people thought in the process of credibility assessment: People firstly judged the credibility of sources, then selected several elements in the information; according to the elements people conducted the assessment and finally got the results; moreover, judging process observed the interest-driving rules, considering efficiency but not accuracy. The judging criterions which had positive effects on assessment included accuracy, authority and objectivity. And the four types of positive affecting cues are information contents, information sources, relevant news of the topic and information spread, every type contained several small cues.

An assistant webpage based on Weibo, helping online users judge the credibility of the news information by visualizing the positive affecting cues in the method, was designed. The first part of the webpage showed the content and source information in Weibo, adding the explanation of terms in the content. And the second part visualizing the news spread information, containing the timeline of the experts' opinion and the cloud of the general users' hot word in the retransmission of the Weibo news. The third part gathered the relevant news in Weibo and added the news in official online news media.

There were three main limitations in the study. Firstly, we had just verified the effectiveness of the method in Weibo news. The initial object of the research was to get a method that could help Internet users for the process of all the online information credibility assessment, and the method was conducted to help people judging online news in all the Internet website, and the webpage was applied only for Weibo. In other words, the assisting range was narrow. Secondly, the webpage only visualized the affecting cues of the method, but without the judging criterions, which were just used to test the assistant effects. At last, the effects of the assistant webpage had not been measured. The future study can concentrate on expanding the applying range of the method, visualizing the judging criterions in the method, and improving the design of the assistant webpage.

References

1. Chaiken, S., Trope, Y.: Dual-Process Theories in Social Psychology. Guilford Press, New York (1999)
2. Rieh, S.Y.: Judgment of information quality and cognitive authority in the web. J. Am. Soc. Inform. Sci. Technol. **53**(2), 145–161 (2002)
3. Fogg, B.J.: Prominence-interpretation theory: explaining how people assess credibility online. In: CHI 2003 Extended Abstracts on Human Factors in Computing Systems, pp. 722–723 (2003)
4. Simon, H.A.: A behavioral model of rational choice. Quart. J. Econ. **69**(1), 99–118 (1955)
5. Hilligoss, B., Rieh, S.Y.: Developing a unifying framework of credibility assessment: construct, heuristics, and interaction in context. Inf. Process. Manage. **44**(4), 1467–1484 (2008)
6. Metzger, M.J., Flanagin, A.J., Medders, R.B.: Social and heuristic approaches to credibility evaluation online. Journal of Communication **60**(3), 413–439 (2010)
7. Metzger, M.J.: Making sense of credibility on the web: models for evaluating online information and recommendations for future research. J. Am. Soc. Inform. Sci. Technol. **58**(13), 2078–2091 (2007)

8. Petty, R.E., Cacioppo, J.T.: The elaboration likelihood model of persuasion. Communication and Persuasion. Central and Peripheral Routes to Attitude Change. Springer Series in Social Psychology, pp. 1–24. Springer, New York (1986)
9. Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A.: The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accid. Anal. Prev. **43**(5), 1666–1676 (2011)
10. Schwarz, J., Morris, M.: Augmenting web pages and search results to support credibility assessment. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1245–1254 (2011)
11. Castillo, C., Mendoza, M., Poblete, B.: Information credibility on twitter. In: Proceedings of the 20th International Conference on World Wide Web, pp. 675–684 (2011)
12. Flanagin, A.J., Metzger, M.J.: The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. New Media Soc. **9**(2), 319–342 (2007)
13. Morris, M. R., Counts, S., Roseway, A., Hoff, A., Schwarz, J.: Tweeting is believing?: Understanding microblog credibility perceptions. In: Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, pp. 441–450 (2012)
14. Armstrong, C.L., McAdams, M.J.: Blogs of information: how gender cues and individual motivations influence perceptions of credibility. J. Comput.-Mediated Commun. **14**(3), 435–456 (2009)
15. Westerman, D., Spence, P.R., Van Der Heide, B.: A social network as information: the effect of system generated reports of connectedness on credibility on twitter. Comput. Hum. Behav. **28**(1), 199–206 (2012)