

# Research on Credibility of Agricultural Information Based on Information Source

Yuyi Zhang <sup>a)</sup>, Yayao Zuo

*Faculty of Computer, Guangdong University of Technology, Guangzhou, 510006 China.*

<sup>a)</sup> Corresponding author: 1540171202@qq.com

**Abstract.** Nowadays, there is a huge amount of webpage information on the Internet, which quality is uneven, and the Internet is full of all kinds of information that is hard to distinguish. In order to solve the credibility problem of agricultural web information, this paper proposes a model to establish the relationship between the information Sources and the information credibility in the agricultural web page information, and quantify the five impact factors affecting agricultural information. Experiments show that this method is feasible and helpful for screening reliable agricultural web information.

**Key words:** agricultural webpage; information source; credibility.

## INTRODUCTION

The study of information credibility is extremely valuable. From the literature [1] we can see that with the increase of people's access to information on the Internet, people's decisions in daily life are more dependent on network information, and the assessment of credibility has become part of people's lives. Regarding the credibility of information, scholars at home and abroad have different definitions for different fields. Domestic scholars Cai Zixing and Xu Guangyou based their efforts on artificial intelligence in the literature [2] and defined the credibility as follows: based on experience, the degree of trust in a thing or phenomenon is true, and gives the uncertainty representation and basic reasoning algorithm. The foreign scholar Iding M, ME Crosby, from the perspective of the creator of trust in literature[3], defines the degree of trust as: the degree of trust or trust. The above explanation of credibility has played a very good guiding role in our research on credibility.

At present, the research on the credibility of information is still in a period of growth, and it is also a hot research field in the Internet age. Scholars at home and abroad have begun to study the credibility of information sources. The literature [4] shows that the credibility of information sources has great impact on the credibility of information on the website. The literature [5] is based on website characteristics, user attributes, and information verification behavior to assess the credibility of the network information. The literature [6] examines the website certificate, the number of advertisements, and the web page design and other factors, so as to achieve an examination of web page information.. From the above, we can see that scholars at home and abroad have studied the properties of websites for information sources. However, few scholars have conducted in-depth research on the reliability of web information in the agricultural sector. In light of this, studying the credibility of agricultural web information is a challenging topic.

This paper has carried on the preliminary exploration to the credibility of the agricultural web page information, constructed the relation model between the information source and the information credibility in the agricultural web page information. By quantifying the five major factors affecting the information source: the credibility of the information publisher; the amount of advertisements on the web pages; the frequency of updating web pages; the ranking of web pages in search engines; the number of web pages visited to calculate the credibility of information sources that affect agricultural information. Experiments show that this method is feasible and helpful for screening reliable agricultural web information.

## THE CREDIBLE MODEL OF AGRICULTURAL WEBPAGE INFORMATION SOURCE

### The choice of information source credibility factor

The website or webpage as a source or carrier of agricultural information directly reflects the credibility of the information. According to the following impact factors, the reliability of the agricultural information webpage can be judged.

1, The credibility of the information publisher, marked as  $f_{cre}$ . The credibility of agricultural web information is closely related to the credibility of the publisher. Publishers of information are generally divided into individuals and organizations. We classify the individuals and organizations of information publishers as follows:

(1) In the individual, the publisher's attributes can be divided into three types: anonymous; net name; real name, and their credibility increases in turn. Among them, the anonymous credibility is the lowest, the real name is the highest, and the net name is the second highest.

(2) Relative to individuals, the credibility of the organization is higher. In the organization, the publisher's attributes can also be divided into three categories: spontaneous group organization, professional organization organization, country or government, and their credibility increases in turn. Or the credibility of the government is the highest, professional organizations are the second, and spontaneous groups are the lowest.

In this paper, we mainly use the questionnaire survey and the AHP approach to delineate the specific impact size of individual and organizational impact factors. The results are shown in Table 1.

**TABLE 1.** Information publisher's weight

Information Publisher	Publisher Properties	fcre
Individual	Anonymous	0.1
	Net name	0.2
	Real name	0.5
Organization	Spontaneous group	0.3
	Professional organization	0.6
	State or government	0.9

2, Webpage ad volume, marked as  $f_{ad}$ . In general, highly reliable web page information will have a correspondingly small amount of advertisements. For example, Guangdong Agricultural Information Network, China Agricultural Information Network, etc, and these agricultural websites are highly trustworthy. Its content is basically related to agricultural information, and there are very few advertisements that have merchants to sell. From this, it can be seen that a high-confidence web page information will have a correspondingly small amount of advertisements on the web page.

3, The frequency of updating web page information, marked as  $f_{fre}$ . The frequency of information update refers to the frequency of updating the webpage content by the webpage publisher. In general, the faster the update frequency, the higher the credibility of the webpage. The timeliness requirements of agricultural information are very high. For example, the price of vegetables, along with changes in supply and demand and weather, will fluctuate in price. It is probable that the price on the first day is 5 yuan. The next day it became 8 yuan. It can be seen that the faster the page information is updated, the higher its credibility will be.

4, The ranking of web pages in search engines, marked as  $f_{rank}$ . In general, the higher the rank of a web page in a search engine, the higher the credibility of its information.

5, Page visits, marked as  $f_{pv}$ , In general, the number of pages visited is proportional to the influence of the page. For example, we want to inquire about the agricultural information of Guangdong Province. Then, we generally go to the Guangdong Agricultural Information Network, and the traffic of this website is very large. As we known, The Guangdong Agricultural Information Network was sponsored by the Guangdong Provincial Department of Agriculture, which can be seen that its influence is also great. The greater the number of pages visited, the higher the credibility of the page.

### Calculation of Information Source Credibility

Based on the above analysis of the indicators affecting the agricultural web page information, there are five main types of information indicators that can affect the credibility of agricultural web page information.

Through the five categories of agricultural web information credibility factors, we define the credibility of information sources as follows:

$$P\_source = \lambda_1 * f_{cre} + \lambda_2 * f_{ad} + \lambda_3 * f_{fre} + \lambda_4 * f_{rank} + \lambda_5 * f_{pv} \quad 1)$$

Among them, for the specific values of the five types of impact factors  $f_{cre}$ ,  $f_{ad}$ ,  $f_{fre}$ ,  $f_{rank}$ ,  $f_{pv}$ , we mainly use third-party tools and the results of questionnaire surveys to translate into relevant values. For index weight  $\lambda_i$ , its value is mainly converted into relevant values through questionnaire survey and expert scoring. As shown in Table 2,  $\lambda_1 > \lambda_3 > \lambda_5 > \lambda_4 > \lambda_2$ , ie, the credibility of the information publisher is the highest, and The weight of the advertisement volume of the webpage is the lowest.

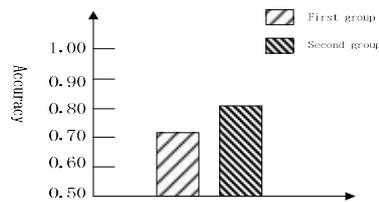
**TABLE 2.** Parameter weight

Weight	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$\lambda_5$
size	0.45	0.05	0.25	0.10	0.15

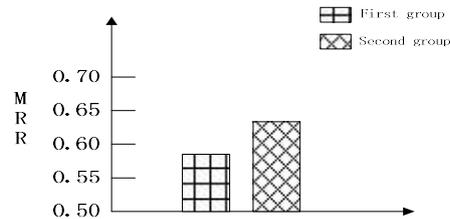
### EXPERIMENT AND ANALYSIS

The experimental platform for this experiment is the linux system Centos 6.5, which is mainly implemented based on python. The data is captured using pyspider. The materials used in the experiment were mainly crawled from the Internet based on agricultural information, including eight agricultural websites, such as: China Agricultural Information Network, China Fishery Information Network, China Fisheries Information Network, etc; and there are also six forums. For example: Aquaculture Technology Forum, Planting Technology Forum, etc. Among them, six hundreds positive cases and two hundreds negative cases were used, with a total of eight hundreds web pages.

The reliability evaluation experiments of web pages in this article are mainly divided into two groups for experiments. The first group is based on the survey website certificate, the number of advertisements, and webpage design factors, which proposed by M. Wassmer and C.M. Eastman, in order to assess the credibility of the webpage information; The second group is the method adopted in this paper, which quantifies the following five factors: credibility of information publishers based on information sources; volume of web advertisements; frequency of web page information updates; ranking of web pages in search engines. The experiment evaluates the credibility of the web page information by using the Mean Reciprocal Rank (MRR) and accuracy values in the retrieval model evaluation method. The result is as follows.



**FIGURE 1.** Accuracy result feedback



**FIGURE 2.** MRR result feedback

From the above figure, we can see from the two angles of accuracy and MRR value that the reliability of the second group is higher. This shows that the information source credibility assessment method proposed in this paper is better.

## CONCLUSION

This paper studies the relationship between web information and information credibility based on agricultural information, and quantifies the five factors that have a greater impact on the credibility of agricultural web information: the credibility of the publisher of the information; the amount of advertisement on the webpage; the frequency of updating the webpage information and the rank of the webpage in the search engine, which can calculate the credibility of information sources that affect agricultural information. Experiments show that this method is feasible and helpful for screening reliable agricultural web information. However, the credibility of information based solely on sources of information is not reliable. By combining the credibility of information content, we can make our information credibility evaluation truly robust. Therefore, the construction of information source and information content credibility model will be the next research direction of this paper.

## ACKNOWLEDGMENTS

At the time of this thesis, I would like to express my deep appreciation to all those who have taken care and help in learning and living during the Master's degree.

First of all, we would like to thank Professor Zuo. Can be successfully completed the writing of the paper, all embodies the teacher's effort and sweat. The teacher in the paper topics, research programs to determine and the specific implementation process have given careful guidance, their rigorous attitude and systematic research ideas I benefit for life.

## REFERENCES

1. N Diviani, BVD Putte, S Giani, JCV Weert. Low Health Literacy and Evaluation of Online Health Information: A Systematic Review of the Literature[J]. *Journal of Medical Internet Research*, 2015, 17(5): e112.
2. Cai Zi-xing, Xu Guang-you. *Artificial Intelligence and Its Applications*[M]. Beijing: Tsinghua University Press, 2005.
3. Iding M, Auernheimer B, Crosby M E. Towards a Metacognitive Approach to Credibility Determination[C]. *Proc Of the 2nd ACM Workshop on Information Credibility on the Web*, 2008: 75-80.
4. AL Ginsca, A Popescu, M Lupu. Credibility in Information Retrieval[J]. *Now Publishers Inc.*, 2015, 9(5): 355-475.
5. E Mitchelstein, PJ Boczkowski. Online News Consumption Research: An Assessment of Past Work and an Agenda for the Future[C]. *New Media & Society*, 2016, 12(7): 1085-1102.
6. M. Wassmer and C. M. Eastman. Automatic evaluation of credibility on the web[J]. *Proceedings of the American Society for Information Science and Technology*, 2010, 42(1): NA-NA.