1-1-2019

Big-Data Based Analysis for Communication Effect of Science-Technology Public Accounts On Social Media

Jinluan Ren
Communication University of China

Wen Cao
China Mobile Group Beijing Co. Ltd

Bo Li
Communication University of China

Lihua Liu
Communication University of China

Lin Cai
Communication University of China

MSU Digital Commons Citation

Authors

Jinluan Ren, Wen Cao, Bo Li, Lihua Liu, Lin Cai, and Ruben Xing

Big-Data Based Analysis for Communication Effect of Science-Technology Public Accounts on Social Media

Jinluan Ren, Communication University of China, Beijing, China

Wen Cao, China Mobile Group Beijing Co., Ltd., Beijing, China

Bo Li, Communication University of China, Beijing, China

Lihua Liu, Communication University of China, Beijing, China

Lin Cai, Communication University of China, Beijing, China

Ruben Xing, Montclair State University, Montclair, USA

ABSTRACT

Public accounts on social media have become important channels for information dissemination. Well-designed public social media accounts are vital to better communicate science and technology (S-T) achievements. This article defines the S-T communication concept and proposes the analyzing dimensions. In order to measure the communication effect, this research collected 7,246 articles from S-T public accounts on WeChat. We analysis these massive data incorporating neural network (NN) and multivariate linear regression (MLR) model. The evaluation indicator system of communication effect includes three levels indicators. The research found the following factors affecting the S-T communication effect in different degrees: the number of active fans on Science Technology Public Accounts on Social Media (STPA-SM), locations where the articles are published, the authentication status of STPA-SM, and so on. Finally, the article proposes some strategic suggestions for improving the communication effects of S-T achievements through STPA-SM.

KEYWORDS

Big Data, Communication Effect, Public Accounts, Social Media

INTRODUCTION

Since the 21st century, the social media environment has become more and more digitized, socialized and mobile (Ren, Cao, Liu, Huang & Zhu, 2017). The mode of S-T communication on the traditional Internet is gradually adapted to the mobile Internet. Therefore, in the era of emerging mobile Internet, the public accounts on social media, such as WeChat, have opened up new channels for S-T communications. Social media has become a major channel of S-T diffusions. However, due

to the high degree of professionalism and rigor, S-T achievements are disseminating by relatively smaller number institutions or individuals compared with content in other fields, thus communication effects are not satisfactory (Zhou, Xu, Wang, Zhu, Luo, Deng &Wu, 2016). Most scientific research institutions and enterprises who have set up STPA-SM are just following the trend of the Internet development, but the actual results are not ideal yet.

Based on the propagation characteristics and operation situation of STPA-SM, this paper proposes a series of hypotheses regarding the influence degree of different factors on communication effects and constructs a measure indicator system that affects the communication effect of STPA-SM. We collect 7,246 articles on STPA-SM from QingBo big data platform and study the influence degree of different factors to communication effect combining techniques with NN and MLR model. Finally, this work provides adaptive operation strategies for STPA-SM to improve its' communication effect in the social media era.

LITERATURE REVIEW

The Patterns, Content and Effects of Social Media Used for S-T Propagation

After the mid-1990s, with the rapid development of the Internet, the carrier of modern S-T communication has developed from the traditional books, magazines, newspapers, radio and television to online digital media. Online communication has become an indispensable part of scientific activities (Walsh, 1996; Walsh & Bayma, 1996). Following the success of email communication (Eisend, 2002), social media is beginning to become a new channel for the spread of science and technology (Su, Scheufele, Bell, Brossard & Xenos, 2017).

Social media has caused a paradigm shift in science communication (Su, Scheufele, Bell, Brossard & Xenos, 2017). The traditional thought of science communication has gradually evolved into the idea of emphasizing contact with science (Lee & Vandyke, 2015). Social network platforms provide a lower-cost access to scientific achievements. Both traditional network and social network play important roles in transforming scientific knowledge, such as newspapers, television, the Internet and so on. Social network also provides a more convenient environment for academic communication (Mondragon, Mondragon &Valencia, 2017; Chang, Kim, Kang, Shim & Ma, 2018; Gu & Widén-Wulff, 2011). More and more Chinese scientists prefer to choose different social network platforms to conduct interdisciplinary collaboration, to improve controllability and advocate public participation (Jia, Wang, Miao & Zhu, 2017). More than 900 samples were collected by the online survey groups and shown that people with higher levels of monitoring satisfaction and social utility motivations tend to pay more attention to scientific news on social media (Ho, Yang, Thanwarani &Chan, 2017).

The display forms of STPA-SM are becoming more and more diversified. STPA-SM had many drawbacks at the beginning, such as weak transmission intensity, limited content richness, simple format of information dissemination, insufficient application functions, etc. However, the topics emerged in STPA-SM are gradually diversified, and more multimedia elements such as pictures, animations and video are applied (Cai, 2017).

There are three main types of STPA-SM include portal S-T channels, professional S-T media and private S-T media. Each type has different content emphases. For example, most content of "Tencent Technology" is about popular information, news reports and news research in the S-T industry. "China Science Daily" concentrates on disseminating scientific knowledge and scientific spirits. "Guokr" focuses on popular science knowledge in medical and health fields (Fan, 2015).

Further, increasing attention has been paid to the research on the communication effects of STPA-SM. In the case study of "Guokr", the factors affecting the communication effects of STPA-SM are the article resource, the theme of science, the main line of science communications and the country where the scientific issues occur (Wang, 2017). Selected articles from 783 STPA-SM, Jin and others found that the authenticated department and multimedia presentation style are significantly impacting

the WCI (i.e. WeChat Communication Index). WCI is used to measure the influence of WeChat public accounts, which is a scientific calculation method and designed by QingBo big data platform. The average number of views and likes, and the published content positioning also influence WCI (Jin, Jiang, Chen & Shen, 2017). In the case of the official public account of Ministry of Science and Technology of the People's Republic of China, it is pointed out that the content of articles has a certain relationship with the number of views of social media, which can make the S-T communication more efficient and effective (Zhang, Xu, Wang, Zhu, Luo, Deng & Wu, 2017). Some researchers studied how to improve the communication effect of science knowledge in the new media environment by taking Nano science as an example (Liang, Su, Yeo, Scheufele, Brossard, Xenos, Nealey, 2014).

Current Research Gaps

The current research on the STPA-SM can be summarized as shown in Figure 1. The paradigm of S-T communication is gradually changing by social media. The social media, such as STPA-SM has become one of important patterns for S-T communication. Different types of STPA-SM focus on different promotion content.

At present, there are relatively systematic studies on social media, S-T communication, communication mode, communication content and communication effect. The existing research points out the main factors affecting the communication effect of STPA-SM. However, there are few researches are on following issues: how to use social media to improve S-T communication effect, how to match communication mode with communication content, how to update the communication mode according to communication effect. The research methods are mostly qualitative. Few researches have attempted to quantify the influence degree. The above gaps are shown by the dotted line in Figure 1. We carry out targeted study on the issues above.

S-T COMMUNICATION AND COMMUNICATION EFFECT

Connotation of S-T Communication

The concept of S-T communication was first proposed by Bernal, which pointed out that it was an inefficient method for scientists to communicate mainly depending on scientific journals. Therefore, it was very necessary to construct a science information service system (Bernal, 2003). This was the original source of S-T communication.

In 2000, Wellcome Trust Fund and UKS-T Office defined the concept of S-T communication, which means the communication among S-T community, media organizations, the public, government and so on. (Zhai[, 2002]). Zhai pointed out that S-T communication is a process of knowledge sharing and diffusing across time and space, the main modes of transmission including professional interaction, S-T education, S-T popularization and technology dissemination (Zhai & Yang, 2002).

Figure 1. Analysis of the research status
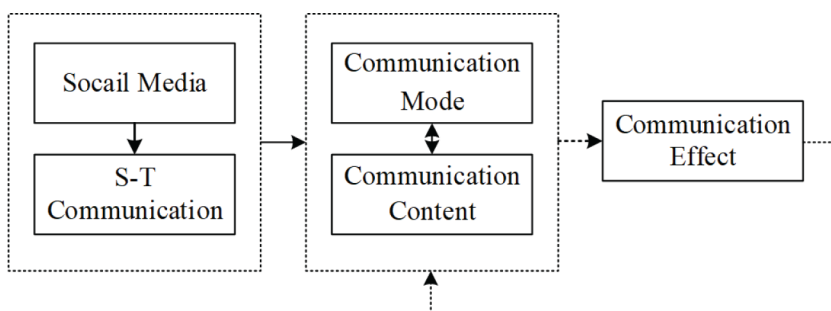
Combining the analysis of S-T communication, this paper regards S-T communication as a diffusion process of S-T knowledge and achievements toward the whole society, which has an impact on the society development.

Communication Effect

The concept of communication effect contains dual meanings. First, it refers to changes in attitudes, psychology, and behaviors that occur when an audience receives a communication behavior with persuasive. Second, it refers to the total effect and consequences of communication activities on the audiences and the society, especially through newspapers, radio, television, and other mass media. (Guo, 1999).

The "social persuasion theory" proposed by Cartwright and Hofland studied the mechanism of communication in detail. In order to achieve effective persuasion, the following two aspects should be done. First of all, it is significant to ensure the information that the communicator wants to spread must reach the sensory organs and consciousness of the recipient. Secondly, the recipient must be informed of the feasibility, necessity and urgency of taking action, and promoted to take appropriate action. There are many factors that will affect the communication effect. Firstly, the communicator should have a high degree of reliability, that is, the identity, status, educational background and life experience of the communicator should match the content to be disseminated. Secondly, the communicator's language expression ability and expression style will also affect the communication effect. Finally, the communication content itself is also very important.

Therefore, according to the previously mentioned analysis, this paper will measure the communication effect of the STPA-SM from three dimensions: the communication content, the communication mode and the communicator.

MEASUREMENT INDICATORS AND INFLUENCING FACTORS OF COMMUNICATION EFFECT OF STPA-SM

Communication Effect Measurement Indicator of STPA-SM

On social media dissemination platform, the views of an article reflect the article's reach, and the number of likes reflects reader's agreement degree of it. Therefore, this paper will use the number of views as an evaluation indicator of spread breadth and use the number of likes as an evaluation indicator of agreement degree to measure the communication effect of STPA-SM. In order to eliminate the influence of the order of magnitude, we take the logarithm of number of views and number of likes respectively. We then perform the weighted average. Assuming the weight of number of views and number of likes are α and $1 - \alpha$ respectively, the communication effect is expressed in Formula 1:

$$Y = \alpha \ln n_1 + (1 - \alpha) \ln n_2 \qquad (1)$$

Here, Y stands for communication effect, n_1 is number of views, which represents the degree of spread breadth, and n_2 is number of likes, which represents the degree of agreement.

Influence Factors of STPA-SM's Communication Effect

According to previous work, this paper measures the communication effect of the STPA-SM from three dimensions: the communication content, the communication mode and the communicator.

The communication content is an important factor affecting the communication effect of STPA-SM. The theme of an article is the concentrated expression of its content. Whether the article

is original or not reflects the novelty of content. Therefore, we constructed the Hypothesis 1 and Hypothesis 2 as follows:

Hypothesis 1: The theme of article published by STPA-SM has a significant impact on the effect of communication.
Hypothesis 2: Whether an article published by STPA-SM is original or not has a significant impact on the effect of communication.

The communication mode could largely affect whether the reader clicks on an article or not. The location of an article may potentially affect the convenience of the reader's click. The publication time provides reading accessibility. And article title is very likely to affect whether it appeals to readers or not. So, the following three assumptions are proposed.

Hypothesis 3: The location of the article published by STPA-SM has significant impact on the effect of communication.
Hypothesis 4: The publishing time of an article on STPA-SM has significant impact on the effect of communication.
Hypothesis 5: The title features of article published on STPA-SM have significant impact on the effect of communication.

The communicator may determine the communication range of articles. The active fans are the basis of communication. The authentication information affects the audience's trust in the public account, so hypothesis 6 and hypothesis 7 are constructed.

Hypothesis 6: The number of active fans of STPA-SM has significant impact on the effect of communication.
Hypothesis 7: The authentication information of STPA-SM has significant impact on the effect of communication.

In order to verify above hypothesis, based on the research of existing literature and analysis of the existing information of STPA-SM, the measurement indicator system of the influencing factors of STPA-SM are designed. Seven secondary indicators and 24 third indicators are designed, as shown in Table 1.

Hypothesis Testing

This paper uses the "List of STPA-SM" released by QingBo Data Platform as data source. QingBo Data Platform has relatively high authority and completeness. According to different authentication type, such as enterprise, scientific research institution, media organization, and individuals, 95 STPA-SM have been selected. The number of articles published in these accounts is about 30,000 in one year. We collected 7,246 articles published from November 19, 2017 to February 20, 2018. The title, published time, location, originality, views number, likes number were obtained.

Using the Boson sentiment analysis engine, the collected social network article titles were analyzed by Python, then the positive and negative sentiment indices of each social network article title were obtained. According to the punctuation marks contained in the title, interrogative and exclamation of the articles' title are marked. According to the subjective judgment, the titles' interesting or scientific are marked. The subject of the article is marked according to the positioning of STPA-SM and the title of the article. So far, we have quantified all the previously designed indicators that might affect the communication effect of STPA-SM, $X = \left(X_1, X_2, ..., X_{24} \right)$.

Table 1. Influence factors of STPA-SM communication effects

Primary Indicators	Secondary Indicators	Third Indicators (Variable Name)	Variable Description
Communication Content	Article Topic	Life science popularization $\left(X_1\right)$	If the theme of the article is life science, its value is 1, otherwise 0.
		News of S-T events $\left(X_2\right)$	If the theme of the article is news of S-T events, its value is 1, otherwise 0.
		Internet hotspots $\left(X_3\right)$	If the theme of the article is Internet hotspots, its value is 1, otherwise 0.
		S-T knowledge $\left(X_4\right)$	If the theme of the article is S-T knowledge, its value is 1, otherwise 0.
		Industrial S-T information $\left(X_5\right)$	If the theme of the article is industrial S-T information, its value is 1, otherwise 0.
	Originality	Whether the article is original $\left(X_6\right)$	If the article is original, its value is 1, otherwise 0.
Communication Mode	Article Location	The location of the article: headline or non-headline $\left(X_7\right)$	If the article location is headline, its value is 1, otherwise 0.
	Published Time	Published time is from 6 to 9($\left(X_8\right)$	If the article was published from 6 o'clock to 9 o'clock, its value is 1, otherwise 0.
		Published time is from 9 to 12 $\left(X_9\right)$	If the article was published from 9 o'clock to 12 o'clock, its value is 1, otherwise 0.
		Published time is from 12 to 14 $\left(X_{10}\right)$	If the article was published from 12 o'clock to 14 o'clock, its value is 1, otherwise 0.
		Published time is from 14 to 17 $\left(X_{11}\right)$	If the article was published from 14 o'clock to 17 o'clock, its value is 1, otherwise 0.
		Published time is from 17 to 19 $\left(X_{12}\right)$	If the article was published from 17 o'clock to 19 o'clock, its value is 1, otherwise 0.
		Published time is from 19 to 22 $\left(X_{13}\right)$	If the article was published from 19 o'clock to 22 o'clock, its value is 1, otherwise0.
		Published time is from 22 to 5 $\left(X_{14}\right)$	If the article was published from 22 o'clock to next 5 o'clock in the morning, its value is 1, otherwise0.
	Title Features	Negative emotional indicator of the article title $\left(X_{15}\right)$	The negative emotional indicator of the title, with a value of 0-1
		Interestingness of the article title $\left(X_{16}\right)$	If the article title is interesting, its value is 1, otherwise 0.
		Scientificity of the article title $\left(X_{17}\right)$	If the article title is scientific, its value is 1, otherwise 0.
		Interrogative of the article title $\left(X_{18}\right)$	If the article title is interrogative, its value is 1, otherwise 0.
		Exclamation of the article title $\left(X_{19}\right)$	If the article title is exclamatory, its value is 1, otherwise 0.

Table 1. Continued

Primary Indicators	Secondary Indicators	Third Indicators (Variable Name)	Variable Description
Communicator	Active Fans Number	Number of active fans of a public account $\left(X_{20}\right)$	Number of active fans of a public account that the article belong to.
	Authentication Information	The authentication type information of a public account that the article belongs to enterprise $\left(X_{21}\right)$	If the public account has been certified as an enterprise, its value is 1, otherwise 0.
		The authentication type information of a public account that the article belongs to scientific research institutions $\left(X_{22}\right)$	If the public account has been certified as a scientific research institutions, its value is 1, otherwise 0.
		The authentication type information of a public account that the article belongs to media organization $\left(X_{23}\right)$	If the public account has been certified as a media organization, its value is 1, otherwise 0.
		There is no authentication type information of a public account that the article belongs to $\left(X_{24}\right)$	If the public account is uncertified, its value is 1, otherwise 0.

In Table 1, except for X_{20} (active fan number), all other variables are all dummy variables. In order to eliminate the influence of the magnitude difference of variables, X_{20} is logarithmically transformed to facilitate modeling. The correlation between each indicator and the effect of communication is analyzed by SPSS. The analysis results are shown in Table 2.

As shown in Table 2, there are significant correlations between article topic, article location, publication time, title feature, originality, active fans number, authentication type and communication effect. Therefore, the above seven hypotheses pass the test.

RESEARCH ON COMMUNICATION EFFECT OF STPA-SM

This paper uses a multi-layer perceptual (MLP) NN and a MLR model to analyze the effect degree of different factors to the communication effect. The former is classic non-parametric classification analysis model and the latter is classic parametric regression analysis model. If the computing results from the two methods are consistent, the reliability of the results is higher. That is to say, if the variables have higher influence degree in both methods then we think they are the more important variables to the communication effect.

Analyzing the Communication Effect Using NN Model

In this paper, we perform a MLP NN analysis. MLP is a feed-forward artificial NN model. In a feed-forward network, information moves only towards one direction, starting from the input layer and then moving through the hidden layer to the output layer. Given a series of features $\mathrm{X} = \left(X_1,\ X_2, ...\right)$ and target Y, MLP can learn the relationship between features and targets for the purpose of classification or regression. According to the indicator system in Table 1, this section selects 24 third indicators in Table 1 as independent variables, and the communication effect as the dependent variable. The model is shown in Equation (2):

$$Y = \alpha lnn_1 + \left(1 - \alpha\right) lnn_2 = f \sum_{k=1}^{24} w_k * X_k + b \tag{2}$$

Table 2. Result of correlation analysis between the impact indicators and communication effects of STPA-SM

Indicators	Indicators/ Correlativity	Communication Effect	Article Topic	Article Location	Publication Time	Title Feature	Originality	Active Fans Number	Authentication Type
Communication Effect	Pearson Correlativity	1	.205**	.233**	.198**	-.074**	.096**	.692**	-.071**
	Significance(Two-Tailed)		.000	.000	.000	.000	.000	.000	.000
Article Topic	Pearson Correlativity	.205**	1	.011	.096**	.029*	.007	.174**	.117**
	Significance(Two-Tailed)	.000		.337	.000	.015	.570	.000	.000
Article Location	Pearson Correlativity	.233**	.011	1	-.134**	-.035**	.181**	-.043**	.040**
	Significance(Two-Tailed)	.000	.337		.000	.003	.000	.000	.001
Publication Time	Pearson Correlativity	.198**	.096**	-.134**	1	-.232**	-.103**	.362**	.111**
	Significance(Two-Tailed)	.000	.000	.000		.000	.000	.000	.000
Title Feature	Pearson Correlativity	-.074**	.029*	-.035**	-.232**	1	-.005	-.105**	-.031**
	Significance(Two-Tailed)	.000	.015	.003	.000		.675	.000	.009
Originality	Pearson Correlativity	.096**	.007	.181**	-.103**	-.005	1	-.075**	-.086**
	Significance(Two-Tailed)	.000	.570	.000	.000	.675		.000	.000
Active Fans Number	Pearson Correlativity	.692**	.174**	-.043**	.362**	-.105**	-.075**	1	-.114**
	Significance(Two-Tailed)	.000	.000	.000	.000	.000	.000		.000
Authentication Type	Pearson Correlativity	-.071**	.117**	.040**	.111**	-.031**	-.086**	-.114**	1
	Significance(Two-Tailed)	.000	.000	.001	.000	.009	.000	.000	

** If Confidence Level (Bilateral) is 0.01, correlativity is significance
* If Confidence Level (Bilateral) is 0.05, correlativity is significance

where $X_k, k = 1,2,\cdots,24$ represents the factors affecting the communication effect, $w_k, k = 1,2,\cdots,24$ is the weight of each input. The main function of deviation b is to provide each node constant value that can be trained outside normal input accepted by each node. f is the activation function. The activation function used in the hidden layer in this model is a hyperbolic tangent function, as shown in Equation (3). The activation function used in the output layer is an identity function:

$$\tan hx = \frac{\sin hx}{\cos hx} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3}$$

It is generally considered that number of views and number of likes are equally important. So set $\alpha = 0.5$ in equation (2). 70% of the samples are taken as the training set, and 30% samples are taken as the test set for MLP modeling. It is determined that the number of hidden layers of the model is 1 and the number of hidden nodes is 4 after training. The prediction accuracy of the model is shown

in Table 3, and the prediction effect is shown in Figure 2 and Figure 3. The higher importance of the independent variable is shown in Figure 4.

The relative error of the training set is 30.1%, while the relative error of the testing set is 24.6%. The scatter plot of the predicted value shows a linear trend. In Figure 3, the residual points are evenly distributed around the straight line of 0, indicating that the NN model has good training effect and degree of fitting.

From Figure 4, using NN analysis method we found that the number of active fans $\left(\ln X_{20} \right)$ of STPA-SM is the most important dependent variable, which has much influence on the communication effect. Followed by whether the article's location is headline $\left(X_7 \right)$, whether the publish time is from 6 o'clock to 9 o'clock $\left(X_8 \right)$, whether it is unauthenticated $\left(X_{24} \right)$, whether it is news of S-T events $\left(X_2 \right)$ and so on. The title of the article has no obvious influence on the communication effect.

Table 3. Communication effect analysis result by NN model

Training	Square sum error	759.478
	Relative error	.301
	Stop rule used	Error is not reduced in one consecutive steps
Testing	Square sum error	17.968
	Relative error	.246

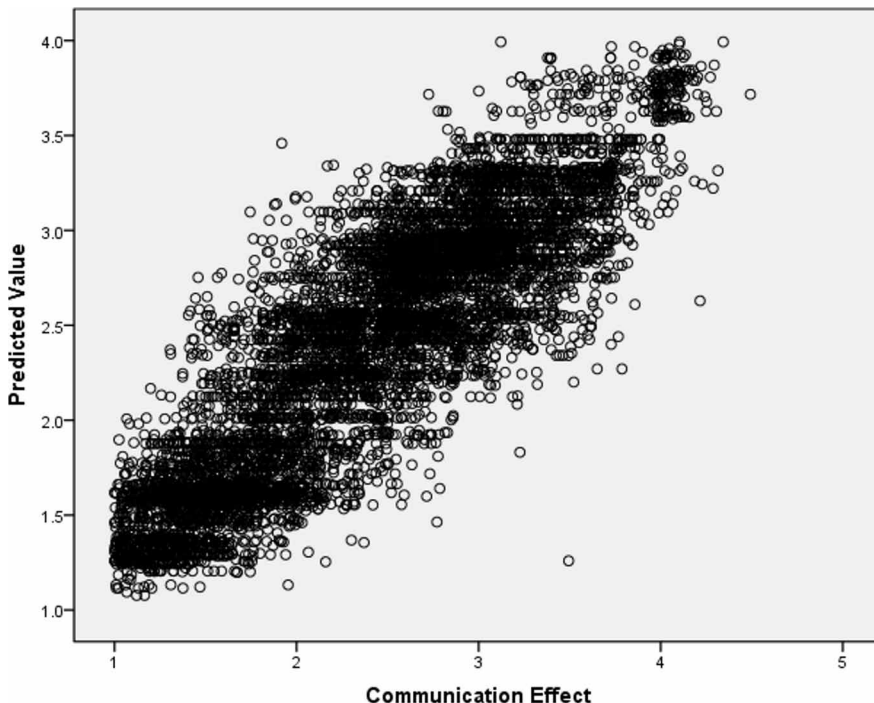**Figure 2. Scatter plot of communication effect prediction value**

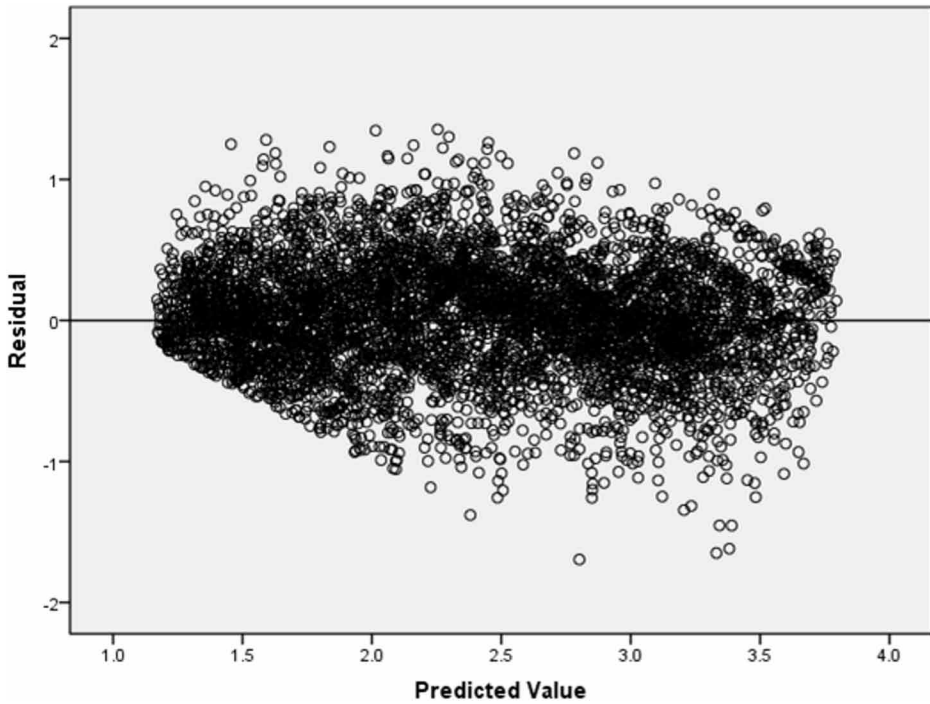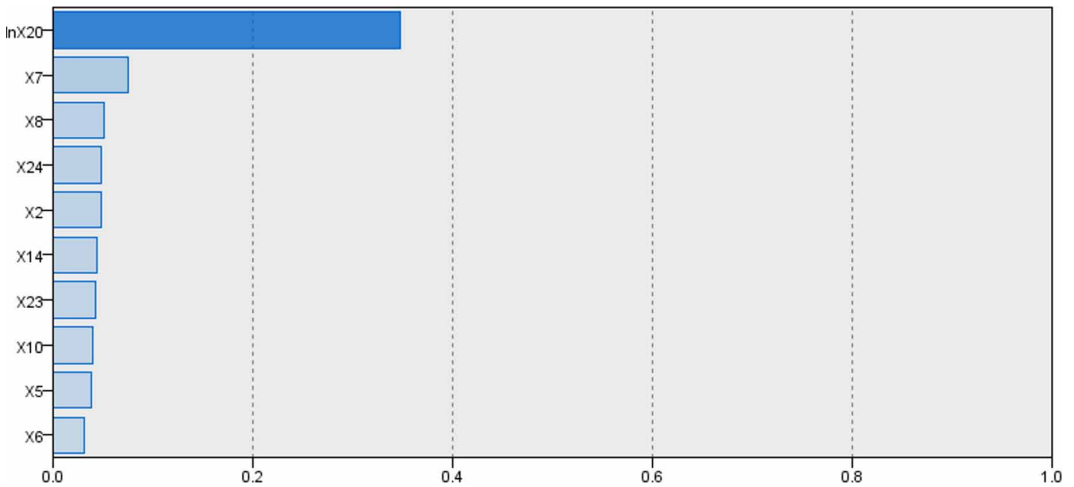**Figure 3. Residual scatter plot of views number prediction value**

Figure 4. The important degree of the main variables affecting the communication effect analyzed by NN model

Analyzing Communication Effect Based on MLR Model

In order to further verify the significance of the influencing factors, this section uses MLR to analyze the communication effect of STPA-SM. We selected 24 three-level indicators in Table 1 as independent variables, and the communication effect as the dependent variable. The model is shown in Equation (4):

$$Y = \alpha lnn_1 + \left(1 - \alpha\right)lnn_2 = \sum_{k=1}^{24}\beta_k X_k + \mu \tag{4}$$

Here $\beta_k \left(k = 1, 2, \cdots, 24\right)$ are the regression coefficients after eliminating the influence of the unit of the dependent variable and the independent variable, indicating the influence degree of the independent variable on the dependent variable. And its value is distributed between 0 to 1. The closer it is to 1, the greater extent to which the independent variable affects the dependent variable. μ is a random variable of the error term, indicating the variability that is contained in Y but cannot be explained by the linear relationship of the k independent variables.

We take $\alpha = 0.5$ and perform regression analysis. Since the number of active fans is significantly larger than other factors, the value of variable is operated by logarithm. From the significance test results of the model, R^2, the coefficient of determination reaches 0.746, indicating that regression model has a good fitness and the independent variable can primely explain the variation of the dependent variable. The statistic variable f is significantly larger than F_α, indicating that the regression model is significant.

From Figure 5, it shows that the most important factor affecting the communication effect analyzed by MLR method is also the number of active fans of the public account $\left(\ln X_{20}\right)$. Followed by whether the article position is headline $\left(X_7\right)$, whether the article is original $\left(X_6\right)$, whether it is unauthenticated $\left(X_{24}\right)$, whether it is authenticated as media $\left(X_{23}\right)$, whether the published time is from 6:00 to 9:00 $\left(X_8\right)$ are all important. These findings are consistent with the previous results obtained from the NN model. Factors including whether it is the negative sentiment index $\left(X_{15}\right)$, whether the topic is Internet hotspot $\left(X_3\right)$, whether the title is interesting $\left(X_{16}\right)$, whether the title is questionable $\left(X_{18}\right)$ are also important, but not appear in the NN model. Whether it is a news of S-T events $\left(X_2\right)$, whether it is industrial S-T information $\left(X_5\right)$, whether it is released from 12 o'clock to 14 o'clock $\left(X_{10}\right)$ are important factors in NN model, however these have not supported by MLR model. Then it indicates that the affecting degree of some factors, such as the characteristics of the title, emotional orientation, Internet hotspot, and news related S-T event and so on needed to be further explored.

Figure 5. The important degree of the main variables affecting the communication effect by MLR model
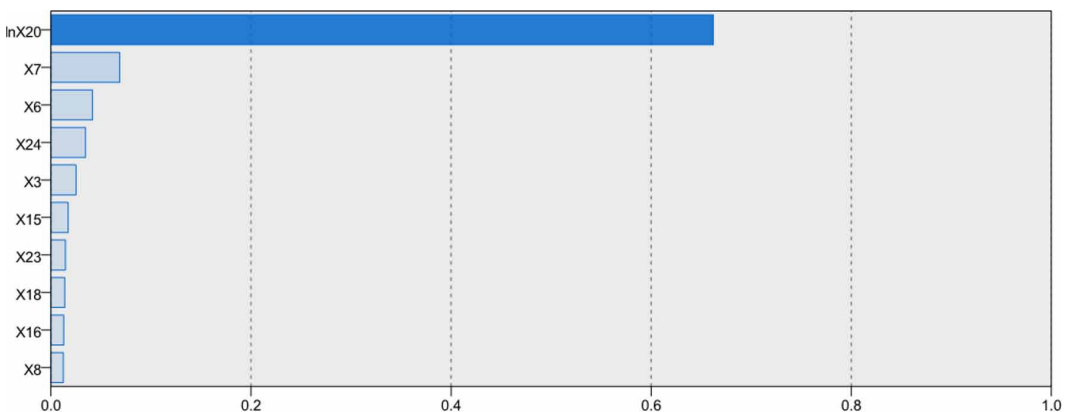
Comparing Results of NN and MLR Method

Comparing the analysis results from NN and MLR method, we found that the main factors and their influence degree of STPA-SM communication effect are mainly consistent. From Figure 4 and Figure 5, the top three influencing factors are same, which are the number of active fans $\left(\ln X_{20}\right)$, whether the article location is headline $\left(X_{7}\right)$, and whether it is unauthenticated $\left(X_{24}\right)$. Hence, the number of active fans has the greatest impact on the communication effects, followed by article location and authenticated information. Whether the published time is from 6 o'clock to 9 o'clock $\left(X_{8}\right)$, whether the article is original $\left(X_{6}\right)$, and whether it is certified as media $\left(X_{23}\right)$ are all considered as important factors although the affecting degree is slightly different. The influence degree of the title features, emotional sentiment, whether it is an Internet hotspot, and whether a news-related S-T event remains need to be further confirmed.

Suggestions

Based on the above research, the following suggestions are proposed on how to improve the communication effect of STPA-SM.

The number of active fans of STPA-SM is found to have the greatest impact on the communication effect, so it is the key to adopt strategies to increase the number of active fans. There will be more readers if an article is published by a STPA-SM with a large number of fans, so the chances of being viewed or liked will increase, then the communication effect will improve. In the early operating stage of STPA-SM, we recommend that it should attract more fans via offline activities and online propaganda. In the medium term, the updating frequency of STPA-SM should maintain in high degree to increase the loyalty of fans. In the long run, the articles on STPA-SM should provide readers with valuable information and knowledge. Then more and more fans will be cling to the public account, and articles on these public accounts have more potential opportunities exposed to a broader population. Finally, the communication effect will be improved.

According to the quantitative analysis results, if the article is released in the headline, the communication effect will be improved. So, it is recommended to place important articles on the headline.

The authoritative certification of STPA-SM does not necessarily have a significant impact on the communication effect. The results of quantitative analysis show that uncertified STPA-SM has achieved better dissemination and recognition. This shows that readers would not have greater recognition of the public number because of authoritative certification information. Therefore, it is recommended that STPA-SM should focus on writing high-quality content in order to improve communication effect.

CONCLUSION

Social media has become an important channel for the S-T achievements communication. This paper collected 7246 public accounts articles on WeChat with the help of data crawling technology. Combining the theory of communication and the characteristics of information dissemination of STPA-SM, we design the evaluation indicator system of communication effect, which is composed of 24 measurement indicators. The influence degrees of various factors on the communication effect are analyzed by MLP NN and MLR method. The following revelation are drawn. Increasing the number of active fans, locating articles in the headline, and improving the quality of the article content play significant roles in improving the communication effect. The interactivity, big data volume and timeliness of social networks provide a more direct way for the dissemination of S-T achievements. It will definitely help to improve the communication effect in the design of STPA-SM referring to the above suggestions.

ACKNOWLEDGMENT

REFERENCES

Cai, Y. (2017) Research on the characteristics of new media science communication: based on content analysis of 6 scientific wechat official accounts. *Studies on Science Popularization. 12*(5), 50-57, 109.

Chang, J., Kim, S., Kang, M., Shim, J., & Ma, D. (2018). The gap in scientific knowledge and role of science communication in South Korea. *Public Understanding of Science (Bristol, England)*, *27*(5), 578–593. doi:10.1177/0963662516685487 PMID:28073306

Eisend, M. (2002). The internet as a new medium for the sciences? the effects of internet use on traditional scientific communication media among social scientists in Germany. *Online Information Review*, *26*(5), 307–317. doi:10.1108/14684520210447877

Fan, X. (2015) The research on communication on science and technology wechat official platform. *Hunan University.*

Gu, F., & Widén-Wulff, G. (2011). Scholarly communication and possible changes in the context of social media: A Finnish case study. *The Electronic Library*, *29*(6), 762–776. doi:10.1108/02640471111187999

Guo, Q. (1999). *Course of communication*. Beijing: China Renmin University Press.

Ho, S. S., Yang, X., Thanwarani, A., & Chan, J. M. (2017). Examining public acquisition of science knowledge from social media in Singapore: An extension of the cognitive mediation model. *Asian Journal of Communication*, *2*(27), 193–212. doi:10.1080/01292986.2016.1240819

Jia, H. P., Wang, D., Miao, W., & Zhu, H. (2017). Encountered but not engaged: Examining the use of social media for science communication by Chinese scientists. *Science Communication*, *39*(5), 646–672. doi:10.1177/1075547017735114

Jin, J., Jiang, S., Chen, A. & Shen, Y. (2017) science communication effects on new media platforms: based on the case of wechat. *Journal of China University of Geosciences,* (2), 107-119.

Lee, N. M., & Vandyke, M. S. (2015). Set it and forget it: The one-way use of social media by government agencies communicating science. *Science Communication*, *37*(4), 533–541. doi:10.1177/1075547015588600

Liang, X., Su, L. Y., Yeo, S. K., Scheufele, D. A., Brossard, D., Xenos, M., & Corley, E. A. et al. (2014). Building buzz: (scientists) communicating science in new media environments. *Journalism & Mass Communication Quarterly*, *91*(4), 772–791. doi:10.1177/1077699014550092

Mondragon, N. I., de Mondragon, L. G., & Valencia, J. (2017). Ebola in the public sphere: A comparison be-tween mass media and social networks. *Science Communication*, *39*(1), 101–124. doi:10.1177/1075547016688908

Ren, J., Cao, W., Liu, L., Huang, R., & Zhu, J. (2017). Analysis of development situation of intelligent media based on technology and market perspective. *Modern Communication*, *39*(10), 133–137.

Su, L. Y., Scheufele, D. A., Bell, L., Brossard, D., & Xenos, M. A. (2017). Information-sharing and community-building: Exploring the use of twitter in science public relations. *Science Communication*, *39*(5), 569–597. doi:10.1177/1075547017734226

Walsh, J. P. (1996). The virtual college: Computer-mediated communication and scientific work. *The Information Society*, *12*(4), 343–363. doi:10.1080/019722496129341

Walsh, J. P., & Bayma, T. (1996). Computer networks and scientific work. *Social Studies of Science*, *26*(3), 661–703. doi:10.1177/030631296026003006

Wang, L., Chen, X. (2017) We-media science communication impact on media salience: an empirical study on "Guokr" wechat public account. *Journalism Bimonthly,* (5), 68-76, 149.

Zhai, J. (2002). *Let technology span time and space*. Beijing Institute of Technology Press.

Zhai, J. & Yang, Z. (2002) Views of concepts of "science communication." *Journal of Beijing Institute of Technology (Social Sciences Edition),* (3), 86-90.

Zhang, H., Xu, S., Wang, Z., Zhu, L., Luo, H., Deng, D., & Wu, H. (2017) Text and data mining of social media in science and technology publicity. proc. In *Portland International Conference on Management of Engineering & Technology* (p. 7).

Zhou, R., Han, F. & Wang, G. (2016) A research on the status and influence of wechat communication of scientific achievements: based on ten wechat official platforms in China. *Studies on Science Popularization, 11*(1), 33-40, 97.

Jinluan Ren is a Professor at the Communication University of China School of Economics and Management. She received her Ph.D degree in Management Science and Engineering from Tianjin University. Research interests in: Media Management and Innovation Management.

Wen Cao received her master's degree in management science and engineering. She is currently an Account Manager at China Mobile Communications Group Beijing Co., Ltd.

Bo Li is an Associate Professor at the Communication University of China School of Data Science and Intelligent Media. She received her Ph.D. degree from Beihang University. Research interests in: applied statistics.

Lihua Liu is a Professor at the Communication University of China School of Economics and Management. She received her master's degree in computer science from Central South University. Research interests in: Information Management, Information Systems.

Lin Cai is a doctoral student at the Communication University of China School of Economics and Management. HResearch interests in media economics.

Ruben Xing is a Professor of School of Management at Montclair State University. He received his EdD degree from Columbia University. Research interests are in information management and business analytics.