Winter 12-6-2018

Research on User Information Behavior and Hotspot Prediction of WeChat Official Accounts Based on Sootoo Network

Shuyu Du
Renmin University of China, Beijing, dududmm@qq.com

Bo Yang
Renmin University of China, Beijing, yangbo_ruc@126.com

Wenjing Chen
Renmin University of China, Beijing, chenwenjing_ruc@163.com

Pengfei Xie
Renmin University of China, Beijing, xiepf94@126.com

Hao Dong
Renmin University of China, Beijing, dh89415@gmail.com

Du, S.Y., Yang, B. Chen, W.J., Xie, P.F., & Dong, H. (2018). Research on user information behavior and hotspot prediction of WeChat official accounts based on Sootoo network. In *Proceedings of The 18th International Conference on Electronic Business* (pp. 557-569). ICEB, Guilin, China, December 2-6.

Research on User Information Behavior and Hotspot Prediction of WeChat Official Accounts Based on Sootoo Network

(Full Paper)

Shuyu Du, Renmin University of China, Beijing, dududmm@qq.com
Bo Yang*, Renmin University of China, Beijing, yangbo_ruc@126.com
Wenjign Chen, Renmin University of China, Beijing, chenwenjing_ruc@163.com
Pengfei Xie, Renmin University of China, Beijing, xiepf94@126.com
Hao Dong Renmin University of China, Beijing, dh89415@gmail.com

ABSTRACT

With the rapid development of mobile Internet, more and more attention has been paid to the new form of online social media represented by WeChat official accounts. However, these operators do not have clear ideas for enhancing WeChat heat, and clear and effective guidance on how to positively influence users' information behavior. Therefore, this paper combines the psychology neutral Stimulus-Response theory with the user information behavior research to construct a second-order user information behavior model. Then, through the official accounts data provided by the Sootoo Network, we collected a total of more than 10,000 push data of nine accounts from January to December, 2017. Based on the empirical research, we find the main influencing factors of user information behavior are identified, and construct a push hotspot prediction model. Also, this paper provides account operator with practical guidance significance in enhancing user information behavior.

Keywords: WeChat official account, user information behavior, hotspot prediction.

*Corresponding author.

INTRODUCTION

WeChat is an instant messaging software based on the mobile Internet launched by Tencent in January 2011. According to the 2017 Ecological Research Report on WeChat users, by the end of 2016, there were more than 10 million official accounts emerging in WeChat (Fang& Shi& Zhang, 2013). Everyone can be a communicator of WeChat we-media platform, but the huge differences in operational capabilities have led to a gradual disintegration of communication power among the communicators.

The general process of users' activities on WeChat official accounts is as follows: firstly, the user receives the push message directly or indirectly. Direct acceptance means that the user receives push messages from the message list. And indirect acceptance means that user receives push messages from other channels such as moments and other users. Secondly, when the user receives the push stimulus, the user can choose to click to read or not to click to read. Then, when the user chooses to read, he receives the stimuli from push text. After that, the user will perform likes, comments, collections, and shares. Finally, after the overall push stimuli, the user will choose follow this account or not (Fang& Lu, 2016).

This paper focuses on the impact of the WeChat official account push on user information behavior, including the reading rate, sharing rate, and new fans. As for the data collection, this paper chose 9 official accounts from which had a huge number of fans accounts data provided by the Sootoo Network, and collected a total of 10,507 push data of nine accounts from January to December, 2017.

The content of this paper includes the following aspects:

(1) Build the user information behavior model based on the Stimulation-Reaction theory in the official account, and summarizes the different influencing factors of each behavioral stage.
(2) Data preprocess: Export the relevant data from the database of official accounts, and then clean, label, normalize, and merge the data. Combining the theoretical model of this paper, define the relevant independent variables and dependent variables.
(3) Data analyze: Use SPSS software to analyze the data and compare the influence of different account types, different push dates, and whether or not to push index on user information behavior.
(4) Build the hotspot prediction model: The hotspots are divided into multiple dimensions: reading rate, sharing rate, comment/collection volume, the number of fans. For different prediction dimensions, build the prediction model based on BP Neural Network and Random Forest.

LITERATURE REVIEW

2.1 User Information Behavior

Davenpor (1997) first proposed the concept of information behavior in 1997. He believed that information behavior is an individual behavior to obtain and process information, including information retrieval, correction, utilization, sharing, storage, and neglect. Wilson (2000) complemented this concept in 2000 and proposed that information behavior includes actively information seeking and utilization, as well as passive information reception. Bai Haiyan (2002) found that user information behavior is a response to the external environment under the control of cognitive thinking. It was based on the information needs and ideological motivations. It went through a process of searching, selecting, and collecting information. Deng Xiaoyong & Li Xiaohong (2008) defined the information behavior of network users as the users in the network environment using network tools to do some activities such as the information retrieval, selecting, exchange, and release under the control of their information needs. Koh J & Kim Y.G (2004) found that the information sharing behavior of virtual community users can drive information sharing of other users, thereby achieving a virtuous circle.

According to the definition of user information behavior, this paper use the user information behavior in the operation of WeChat official accounts as the user's behaviors such as reading, sharing, collection and like for the push from the new media official account (Park J H& Gu B& Leng, 2014).

2.2 Stimulation-Reaction Model

The Stimulation-Reaction Model was first established by John Watson who was the founder of behavioral psychology in the early 20th century. S-R theory pointed out that human's complex behavior could be divided into two parts: stimulus and reaction. Based on John Watson's S-R theory, Howard & Sheth (1969) proposed the consumer's stimulus-body-reaction model. This model believed that the stimuli from the corporate social environment could affect the user's perception and learning, so that the user had a certain degree of emotion and cognition, thus affecting the purchase behavior of consumers. Ergolu (2003) completed the model above through the empirical research, and obtained a new conceptual model. He pointed out that the website atmosphere influenced consumer's pleasure, arousal and attitude, and then influenced consumer satisfaction and approach or avoidance behavior.

In the process of a push from WeChat official account, we define the official account and the push as stimulus. The stimulus mainly comes from the account itself, time attribute, push title attribute, etc before the user accepts information (Wang Ye, 2015). After the user receives the information, the stimulus mainly comes from the text of the push.

2.3 Content Sharing and Communication on Online Social Platforms

Wu Zhongtang (2015) studied the influence of keyword and title on information dissemination through empirical analysis. Fang Xingdong (2013) studied the impact of push frequency, push time, word count of the title, topic on information dissemination and the push’s life cycle through empirical analysis. Yan Yinwen (2017) used BP Neural Network to evaluate the effect of information communication in governmental WeChat official account. Wang Xiudan (2015) studied the influence mechanism of the company’s WeChat official account on users’ willingness to share, and believed that information quality, information volume, information content, frequency of push, and information innovation indirectly affect users’ willingness to share. Yang.J & Counts.S (2010) found that the information content and the reference rate of related users are the key factors affecting the information dissemination.

2.4 Innovation

As for the research variables, this paper broadened the data sources to a certain extent. We obtained the second-hand data of WeChat official account through cooperation with enterpris, which made up for the fact that the variables in previous research were difficult to obtain. Due to data limited, previous researches related to the WeChat official account are currently focused on WeChat marketing strategies, consumer purchase intentions, customer relationship management and so on, but there are few studies on information dissemination and user information behavior. And often use qualitative methods such as questionnaires. This paper uses quantitative empirical analysis to study the influencing factors of user information behavior.

Second, about the theoretical model, based on the theory of user information behavior, this paper combines the Stimulation-Reaction model and the environmental characteristics of WeChat official account. Defines the two-level research concept of push stimuli-user information behavior, and conducts specific dimensionality of user information behavior in this environment. At the same time, find out the main factors affecting user information behavior.

DATA COLLECTION AND PREPROCESSING

3.1 Data Collection

The data in this paper is derived from several official accounts operated by Sootoo Netwok. This paper selects three types of representative official accounts categories: "emotional psychology", "video entertainment" and "science & technology finance". Each category selects 3 accounts with lots of fans to research. This paper has obtained all the push data of these 9 accounts from

January 1 to December 31, 2017, with a total of 10,524 pushes amount, a total of 250 million reading amount, a total of 4,905,000 sharing amount, and a total of 1,829,000 fans growth. The amount of data is shown in Table 1.

Table 1: Data overview

Name	Push amount	Fans amount	New fans amount	Reading amount	Sharing amount	Likes amount
Short Novel	1,133	518,352	93,391	16,752,141	286,425	198,059
Psychological Test	1,164	1,273,403	164,800	44,224,103	847,806	128,685
Mood Signature	1,291	1,036,347	371,340	37,743,509	784,482	263,176
Creative Agency	1,410	1,246,621	250,982	53,010,159	1,447,299	346,485
Movie Heaven	1,112	1,401,068	598,824	26,508,795	236,465	128,611
Miao is coming	2,056	504,679	233,561	43,380,478	907,937	1,381,942
Electric Business News	527	255,097	10,074	5,212,894	58,055	11,202
Internet New Things	1,037	404,068	78,002	14,876,471	202,129	44,278
Financial Reference	794	234,827	27,836	9,942,651	134,625	35,677
TOTAL	10,524	6,874,462	1,828,810	251,651,201	4,905,223	2,538,115

After an in-depth analysis of the model and the variables, the first-order stimulation selected in this paper include the attributes of the account and the attributes of the title. The second- order stimulation selected the attributes of the content, and extract relevant data dimensions from the original data, including the name of WeChat, the number of successful pushes, total number of fans, whether is copyright, number of likes, total readings, number of reprints, title, number of pictures, whether is push index, location, number of characters, date of push, graphic page, number of viewers of the graphic page, number of shares, number of collectors, total number of graphic readers, total number of graphic readings, number of public session readers, number of public session readers, history Message page readers, historical message page readers, moments readers, friends share readers, other scene readers, public number of conversations forwarded friends, public Number of sessions, number of friends forwarded by friends, number of friends forwarded through friends, number of friends forwarded through friends, number of friends forwarded through other scenes, and number of times friends forwarded by other scenes (Wang Xiaoli, 2015). Each field is standardized into an Excel to prepare for subsequent data analysis.

3.2 Variable Definitions

Table 2: Variable definition details

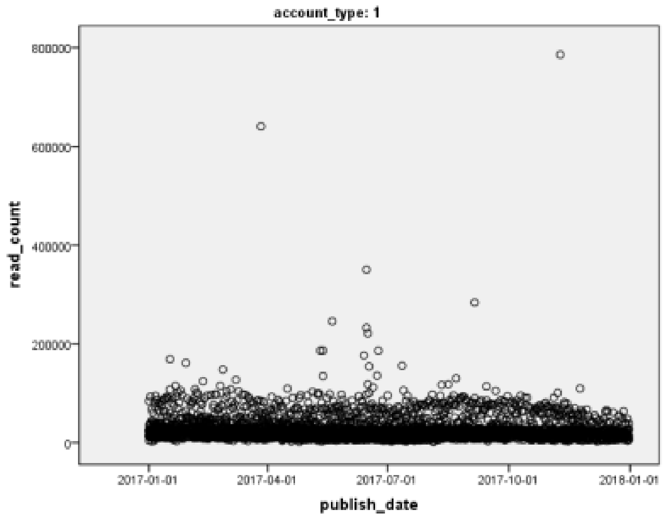
Independent variable	The name of the variable	Meaning	Type
the attributes of the account	wx_officalaccount_name	The name of WeChat	String
	account_type	The type of the official account	Value（N）
	fans	Total of fans	Value（N）
	is_weekend	Whether is weekend	0-1
the attributes of the title	is_question	Whether is a question	0-1
	is_exclamation	Whether is an exclamation	0-1
	is_doublehead	Whether is a doublehead	0-1

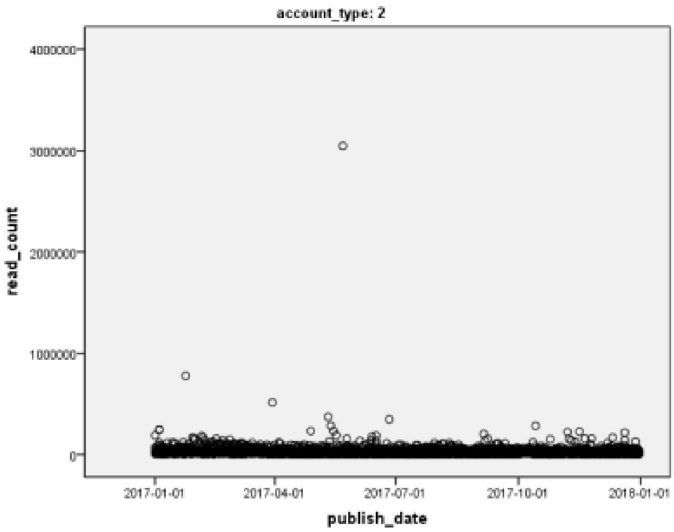
	have_number	Whether has numbers	0-1
	is_first	Whether is the first	0-1
	idx	The position of the push	Value（N）
	title_num	The number of the title	Value（N）
the attributes of the content	img_num	The number of the image	Value（N）
	copyright_type	Whether is copyright	0-1
	article_word	The word of the article	Value（N）
Dependent variable	The name of the variable	Meaning	Type
New fans	new_fans	The number of the new fans	Value（N）
	new_fans_rate	The rate of the new fans	Value（N）
Reading	read_count	The number of reading	Value（N）
	read_count_rate	The rate of reading	Value（N）
	int_page_from_session_read_count	Reading times in account	Value（N）
	1stread_count_rate	The rate of reading one time	Value（N）
	int_page_from_feed_read_count	Reading times in moments	Value（N）
	int_page_from_friends_read_count	Reading times in sharing	Value（N）
Sharing	share_count	The number of sharing	Value（N）
	share_count_rate	The rate of sharing	Value（N）
	int_page_read_user	The reading users in graphic page	Value（N）
	int_page_read_count	The reading times in graphic page	Value（N）
	share_user	The sharing users	Value（N）
	feed_share_from_session_count	The number of sharing from account	Value（N）
	1stshare_count_rate	The rate of sharing one time	Value（N）
	feed_share_from_feed_cnt	The number of sharing from moments	Value（N）
Collect Likes	add_to_fav_count	The number of collecting	Value（N）
	like_num	The number of likes	Value（N）

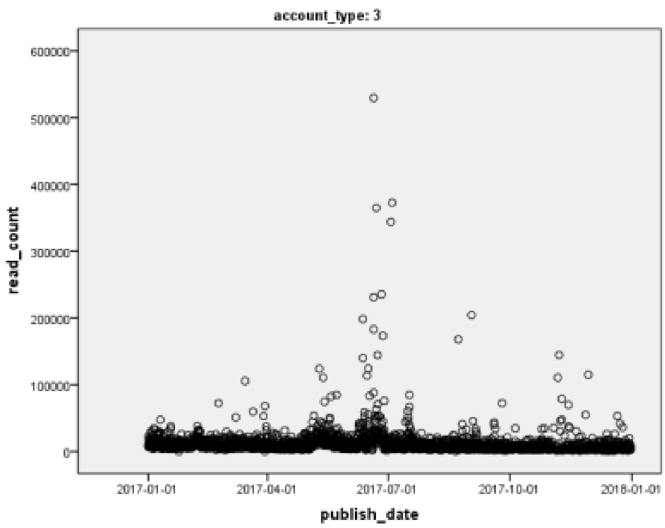
3.3 Outlier Removal

There may be abnormal data in the data collected in this paper, such as “bursting text” which has a high rate of reading and sharing. Its occurrence has great contingency and external reasons. Therefore, in the study of this paper, abnormal data should be removed to find the universal law (Jansen B J, 2000).

The scatter plots of the three types of the official accounts are shown in the figure:


account_type: 1
read_count
800000
600000
400000
200000
0
2017-01-01
2017-04-01
2017-07-01
2017-10-01
2018-01-01
publish_date


account_type: 2
read_count
4000000
3000000
2000000
1000000
0
2017-01-01
2017-04-01
2017-07-01
2017-10-01
2018-01-01
publish_date


account_type: 3
read_count
600000
500000
400000
300000
200000
100000
0
2017-01-01
2017-04-01
2017-07-01
2017-10-01
2018-01-01
publish_date

After removing the outliers:

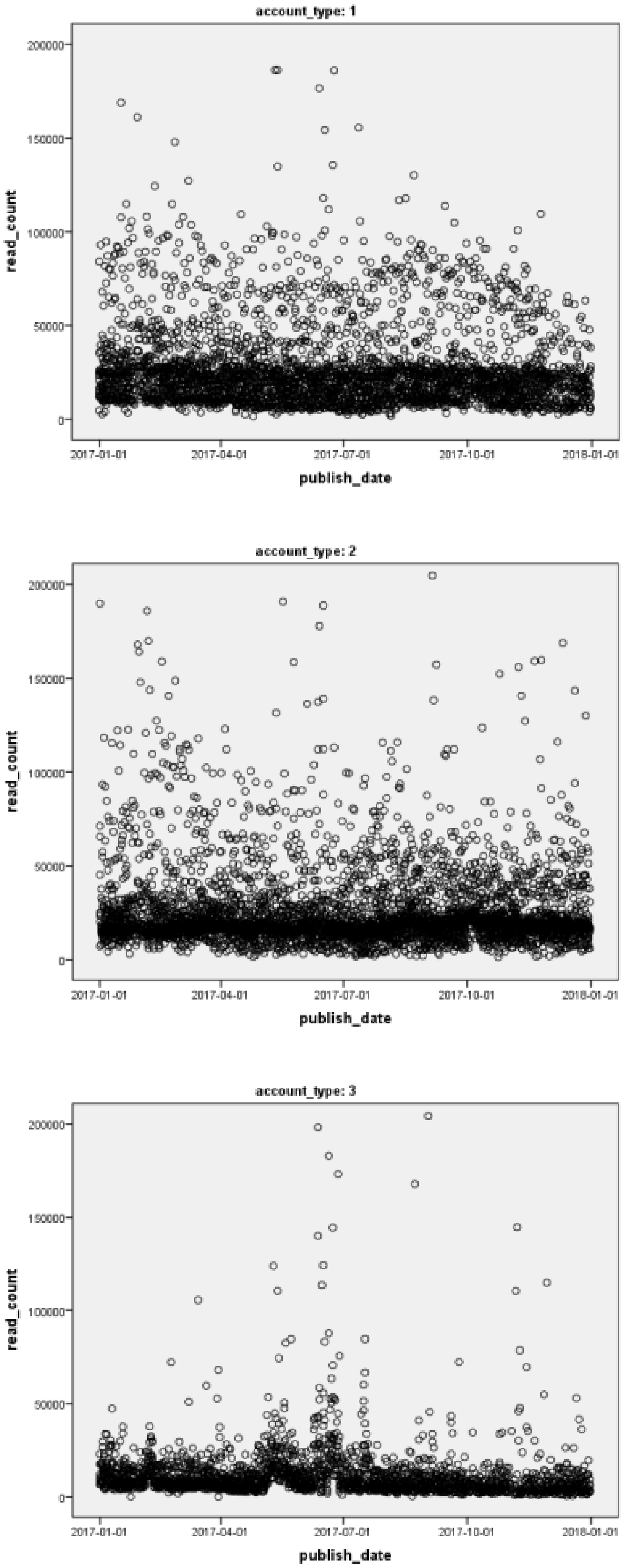
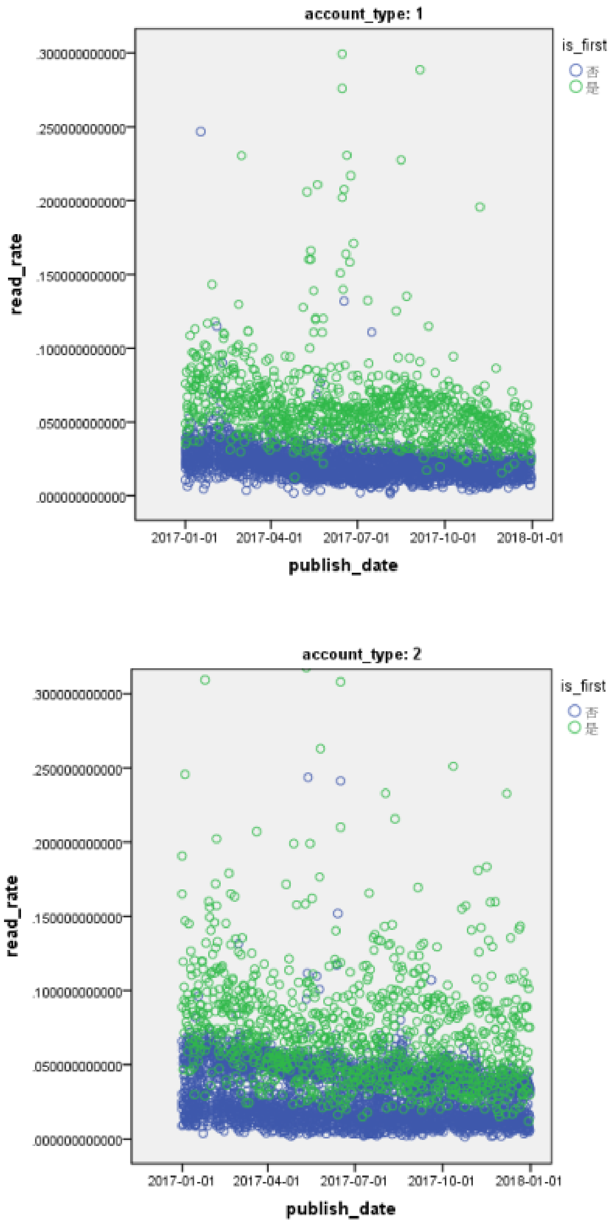


Figure 1: Three types of official account reading date scatter plot

CONSTUCTION AND APPLICATION OF PUSH HOTSPOT PREDICTION MODEL

4.1 Statistical Analysis of Data

Firstly, we analyze the impact of variables on user information behavior, mainly using statistical analysis, one-way ANOVA, correlation analysis and so on. As a result of many and similar variables, we take the effect of the push index on the reading rate as an example, The horizontal axis of the scatter plot is time, the longitudinal axis is the reading rate, the green point is the top push, the blue is not the top push, we can see the obvious difference. As shown in Figure 2.



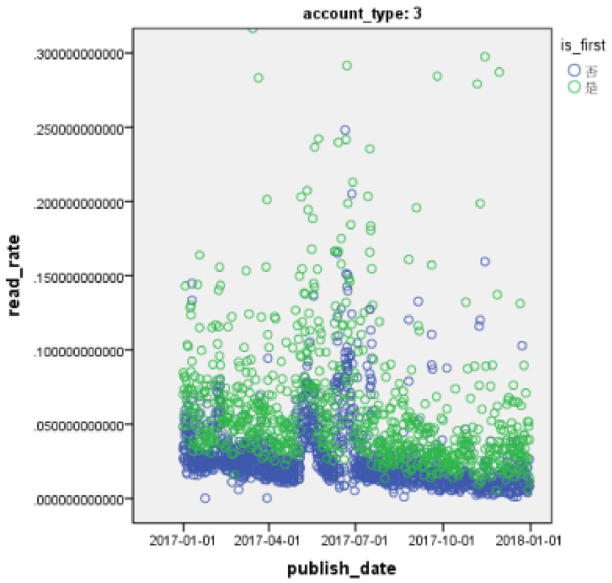
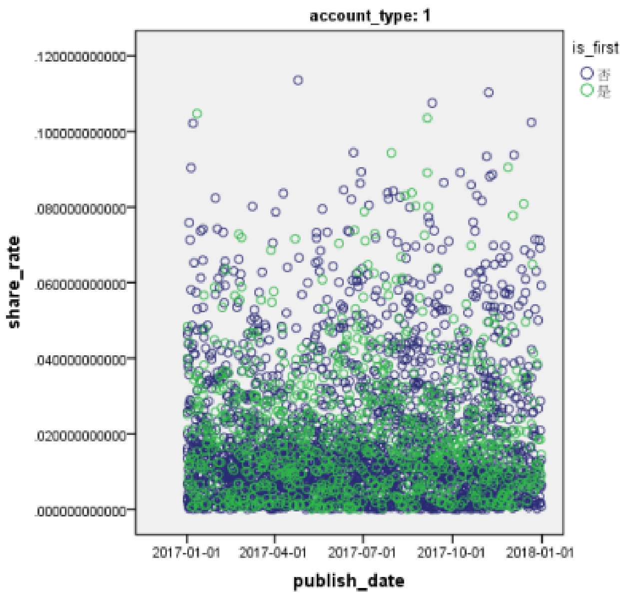


Figure 2: Three type of official accounts reading rate scatter plot

In order to further study whether the push index has a significant impact on the reading rate, we randomly selected 500 samples in the top pushes and non top pushes of each official account to carry out a single factor analysis of variance. The F values of the three groups were 2470,716,172 and the sig values were less than 0.05, reaching a significant level, indicating whether the top pushes had a significant impact on the reading rate under the three types of official accounts.

The sharing is similar, the push index is the first order stimulation, and the effect on the second order reaction sharing is not obvious. It is found that the sig values were less than 0.05 by nonparametric test, so it is significant. The scatter diagram of the influence of the push index on the sharing rate is shown in Figure 3.



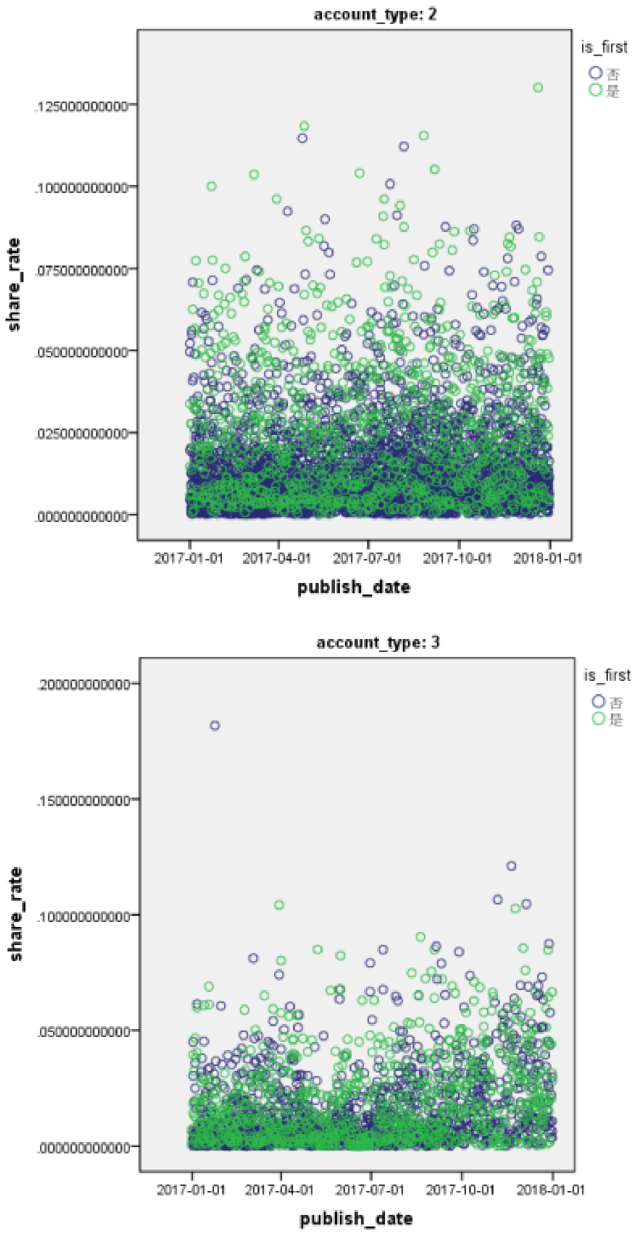


Figure 3: Three type of official accounts sharing rate date scatter plot

After the analysis of the variables, the following conclusions can be obtained as shown in Table 3.

In order to build the prediction model of push hotspots, we define the prediction variable of pushing "hot spots". In this paper, the reading rate and sharing rate in the first 10% intervals of the account is defined as hot push, and then 0-1 categorical variable is constructed. The total variables and the standardized variables related to the reading rate and sharing rate were used as the two set of feature variables. Neural network algorithm and random forest algorithm are used to construct the model respectively.
After data processing, the push for prediction model construction is shown in Table 4.

Table 3: Results of analysis

First order stimulus-first order reaction and second order reaction	**Second order stimulus-second order reaction**
The weekend has a positive impact on the reading rate of "emotional psychology" and "video entertainment", a negative impact on "science & technology finance", and has a positive effect on the sharing rate of all the accounts.	There is a moderate positive correlation between copyright and sharing/ collecting/ likes
In terms of tittles' emotion, interrogative sentences and exclamations have a positive influence on the reading rate of "emotional psychology" and "video entertainment", and have a negative impact on "science & technology finance", and has a positive effect on the sharing rate of all the accounts.	There is a moderate positive correlation between the number of pictures and the sharing rate.
Double tittles has a negative impact on reading and sharing rate, and has a positive effect on increasing the number of fans.	The number of words of the push is moderately positively related to the sharing rate and the amount of collection, and has a negative impact on the growth of fans.
Push index obviously have a strong positive impact on reading rate, and have a moderate positive correlation with the second order reaction (sharing, collecting, and like).	Copyright and push index had no significant impact on the growth of fans, but their interaction had a significant impact on the number of fans.

Table 4: The push for prediction model construction

Variable	Value
Total push number	10507
Hot push number	242

4.2 Experimental Design

The experiment uses 3-fold crossValidation , the data is predicted for three times, and two copies of the three equal data set are taken as the training set, one is used as the test set and the average value is used as the prediction result. We use BP neural network (Li Huanning& Wang Shuning, 2000) and random forest (Zhang Junliang, 2013) two algorithms, each group algorithm uses all variables E1 and correlation variable E2 two groups of variables combination. The two set of feature variables is detailed as follows:

All variables set E1: whether weekends, whether interrogative sentences, whether exclamatory sentences, whether double title , whether including numbers, push index, push positions, copyright, title numbers, number of pictures, the number of words of the push.

Related variables set E2: whether exclamatory sentences, whether double title, push index, copyright, number of pictures, the number of words of the push.

4.3 Modeling Process

Firstly, we model the BP neural network. After a number of experiments, all variables and related variables were selected. The two sets of parameters were used to construct the parameters of the model respectively. The results are shown in the following table 5:

Table 5: Network structure comparison table

	the number of hidden nodes	Training function
all variables	5	Trainlm
related variables	4	Traingd

Then, using random forest modeling (Quinln, 1986), from the original training sample set N, n samples are repeatedly extracted randomly to generate a new training sample set for training decision trees. Then, according to the above steps, m decision trees are generated to form random forests. The classification results of new data are determined according to the scores formed by voting of classification trees.

The code is as follows:

```
results = []
# The parameter value of the minimum leaf node
sample_leaf_options = list(range(1, 500, 3))
# Number parameter value of decision tree
n_estimators_options = list(range(1, 1000, 5))
groud_truth = train_data['Survived'][601:]

for leaf_size in sample_leaf_options:
for n_estimators_size in n_estimators_options:
alg = RandomForestClassifier(min_samples_leaf=leaf_size, n_estimators=n_estimators_size, random_state=50)
        alg.fit(train_data[predictors][:600],train_data['Survived'] [:600]) predict = alg.predict(train_data[predictors][601:])
        results.append((leaf_size, n_estimators_size, (groud_truth == predict).mean()))
print((groud_truth == predict).mean())
print(max(results, key=lambda x: x[2]))
```

4.4 Evaluation Index and Experimental results

We use four statistics in the confusion matrix as the evaluation index of classification algorithm.

Confusion matrix		Prediction outcome	
		N	P
Actural value	N	True negative(TN)	False positive(FP)
	P	False negative(FN)	True positive(TP)

The calculation formulas are as follows:

$$Accuracy = (TP + TN)/(TP + FP + FN + TN) \times 100\%$$

$$Precision = TP/(FP + TP) \times 100\%$$

$$recall = TP/(TP + FN) \times 100\%$$

$$F1 - score = 2 \times precision \times recall/(precision + recall) \times 100\%$$

$$TPR = TP/(TP + FN)$$

$$FNR = FN/(TP + FN)$$

$$FPR = FP/(FP + TN)$$

$$TNR = TN/(TN + FP)$$

The transverse coordinate of the ROC curve is FPR and the longitudinal coordinate is TPR. The curve reflects the sensitivity and specificity of the prediction results as the threshold changes. The larger the area under the ROC curve, the better the performance of the model.

AUC is the graphic area between the ROC curve and the abscissa.

Respectively using the BP neural network algorithm and the random forest algorithm, and characteristic variable groups E1 and E2 were used for comparative experiments. The results of the experimental evaluation are shown in Table 6 as follows.

Table 6: Comparison of experimental results

Model	Variable	Accuracy	Precision	Recall	F1-score	AUC
BP	E1	0.845	0.811	0.293	0.430	0.811
	E2	0.873	0.768	0.467	0.581	0.867
RF	E1	0.845	0.649	0.551	0.596	0.851
	E2	0.861	0.651	0.548	0.595	0.865

By comparing several evaluation indexes, in most cases, the prediction performance of the E2 feature set is better than that of the E1 set. BPNN achieved better accuracy and precision, while RF achieved better recall, F1 score and AUC value, but the superior model of each evaluation index did not gain an overwhelming advantage. However, this paper still chooses BPNN as the better prediction model. The reason is that, when selecting E2 as the feature set, although the F1 score of BPNN and the AUC value are less than RF, the difference is very small. Two is the obvious advantage of BPNN in precision so that it can identify more potential hot spots, which is of more practical significance in the case that hot push is relatively small.

SUMMARY AND PROSPECT

5.1 Summary

This paper based on the stimulus response model in psychological theory, constructs the user information behavior impact model based on the WeChat official accounts, and complements the blank field of WeChat research.

Through a comprehensive analysis, the following conclusions are reached.

1. First order stimulus -- first order reaction and second order reaction

(1) In terms of the date of push, the weekend has a positive impact on the reading rate of "emotional psychology" and "video entertainment". For "science & technology finance", the weekend has a negative impact on the reading rate. The weekend has a positive effect on the sharing rate for all the accounts. This may be due to the nature of accounts. Users will reduce professional reading on non-working days.

(2) In terms of emotion, the appropriate use of interrogative sentences and exclamations can have a positive effect on reading rate of "emotional psychology" and "video entertainment", and exclamations can more positively promote user information behavior than interrogative sentences. This proves the effectiveness of the "title party" to some extent. However, for " science & technology finance ", the behavior of the "title party" will have a negative impact on reading, sharing and fans. For more professional push, users prefer rational and objective description.

(3) Whether double title have a weak negative impact on reading and sharing, and have a weak positive impact on increasing the number of fans.

(4) There is a strong positive correlation between top push and reading rate.

2. Second order stimulation -- second order reaction

(1)There is a moderate positive correlation between push index and second order responses (sharing, collecting, and like).

(2) There is a moderate positive correlation between the number of pictures and the sharing rate, which is not related to the first order response.

(3) Word count of the push is positively related to the rate of sharing and collecting, and is negatively related to the number of fans. The lengthy push may make the reader feel boring, and unrelated to the first order reaction.

(4) Copyright and push index had no significant impact on the growth of fans, but their interaction had a significant impact on the number of fans.

In addition, this paper uses two algorithms and two variables to build a prediction model of user information behavior, and analyzes the effects of different prediction models. It is hoped that the public operators can use more scientific methods and more comprehensive indicators to identify potential push heat.

5.2 Suggestion and Prospect

After integrating data analysis conclusions and forecasting models, this study proposes the following suggestions for WeChat official accounts.

Firstly, reasonable set up the title of the push: strong emotional titles for leisure accounts, professional accounts should be as objective as possible; use a series of content to attract new fans, keep old fans. Secondly, make full use of the top push exposure advantage, place high quality original article, make full use of the interaction between top push and original, so as to get higher sharing rate and number of fans. Then content quality is still the most important factor in communication. Although the "title party" can improve the reading rate to a certain extent, its influence on the second order reaction is usually very small, only by winning the text content can effectively improve the user's second order information behavior.

The shortcomings and expectations of this study are as follows:

The data dimension is limited. In this paper, 3 types of accounts are selected for analysis, it is impossible to compare horizontally between different accounts. Moreover, the push time of the same account is basically fixed, so it is difficult to study the impact of push time.

The text content was not included in the stimulus variable. The two order stimulus of this study only considered the macro variables of each push, and did not divide the important content of text analysis more carefully.

It is difficult to measure the number of fans in user information behavior. Since the background data can only be extracted to each account daily, and cannot refine the new number of fans to each push, so this study does not consider the difference between the fans of the same account in one day.

In the stimulus response model, the body's arousal is actually a part of it, but because this study can't be analyzed by the body arose qualitatively and quantitatively, it is not considered. In the follow-up study, we can consider the introduction of questionnaires and interviews for analysis.

ACKNOWLEDGMENT

Financial support from the Science and Technology Plan Project of Beijing (No. Z171100000117009) and NSFC (No. 91546125) is acknowledged.

REFERENCES

[1] Bai Haiyan, Zhao Lihui. (2002). The Behavior Analysis of Users Information under the Internet Conditions. *Journal of Yanshan University (Philosophy and Social Science Edition), 2002*(2), 87-89. (in Chinese)
[2] Davenport T. (1997). *Information ecology* (83-84). Oxford: Oxford University Press.
[3] Deng Xiaoyong, Li Xiaohong. (2008). Analysis of User Information Behavior in Network Environment. *Information Science, 2008*(12), 1810-1813. (in Chinese)
[4] Eroglu S A, Machleit K A, Davis L M. (2003). Empirical testing of a model of online store atmospherics and shopper responses. *Psychology & Marketing, 20* (2), 139-150.
[5] Fang Jing, Lu Wei. (2016). A Study on Influential Factors of Wechat Public Accounts Information Transmission Hotness. *Journal Of Information, 35*(2): 157-162. (in Chinese)
[6] Fang Xingdong, Shi Xiansheng, Zhang Xiaorong. (2013). Research on WeChat Communication Mechanism and Governance[. *Modern Communication, 2013*(6), 122 -127. (in Chinese)
[7] Fang Xingdong, Shi Xiansheng,Zhang Xiaorong. (2013). Research on the communication mechanism and governance of WeChat. *Modern Communication,2013*(6), 122 -127． (in Chinese)
[8] Howard J， Sheth J N. (1969). *The theory of buyer behavior*. New York: Wiley.
[9] Jansen B J, et al. (2000). Real life, real user, real need: a study and analysis of user queries on the Web. *Information Processing and Management, 36*(2), 207-227.
[10] Koh J, Kim Y G. (2004). Knowledge Sharing in Virtual Communities: an E-Business Perspective. *Expert Systems With Application, 2004*(26), 155-166.
[11] Li Huanning, Wang Shuning. (2000). An Improved BP Neural Network Prediction Method and Its Application. *Systems Engineering, 18*(5), 76-78. (in Chinese)
[12] Park J H, Gu B, Leng A C M, et al. (2014). An investigation of information sharing and seeking behavior in online investment communities. *Computers in Human Behavior, 2014*(31), 1-12.
[13] Quinlna J R. (1986). Induction of Decision Trees. *Machine Learning, 1*, 81-106.
[14] Wang Xiaoli. (2015). Simulation Research on WeChat Information Diffusion Based on Intelligent Multi-agent Networks. *New Technology of Library and Information Service, 31*(6), 85-92． (in Chinese)
[15] Wang Xiudan. (2015). The Research on the Impact of Wechat Marketing on Consumers' Sharing Intention. (Master's thesis, Hangzhou Dianzi University, Hangzhou, in Chinese).
[16] Wang Ye. (2015). Comparative Analysis of the Differences between Weibo and WeChat Marketing Communication Modes. *Public Communication of Science & Technology, 2015*(7). (in Chinese)
[17] Wilson T D. (2000) Human information behavior. *Informing science, 3*(2), 49-56.
[18] Wu Zhongtang, Liu Jianhui, Tang Zhenhua. (2015). Research on the Influencing Factors of the Information Communication in WeChat Public. *Information Science, 2015*(4), 122- 126. (in Chinese)
[19] Yan Yiwen, Zhang Haitao, Sun Siyang. (2017). Research on Evaluation of Information Transmission Effect of Government WeChat Public Based on BP Neural Network. *Library and Information Service, 61*(20), 53-62. (in Chinese)
[20] Yang J, Counts S. (2010). Predicting the Speed, Scale, and Range of Information Diffusion in Twitter. *ICWSM, 2010* (10), 355-358.
[21] Zhang Junliang. (2013). Neural network from entry to mastery. Beijing: Mechanical industry press.
(*Full reference list is available upon request from the corresponding author.)