An investigation on the Factors Influencing the dissemination of WeChat Push Based on HSM and the Prediction of its Content Hotspot

Wenjing Chen
Remin University of China
chenwenjing_ruc@163.com

Bo Yang*
Remin University of China
yangbo_ruc@126.com

Pengfei Xie
Remin University of China
xiepf94@126.com

Huayang She
Remin University of China
shehuayang@ruc.edu.cn

Kanliang Wang
Remin University of China
kanliang.wang@gmail.com

Abstract

With the continuous development of information technology, the carrier of we-media has emerged. The WeChat Subscription Accounts has quickly led the other we-media platforms. During the six years of its emergence, WeChat Subscription Accounts have attracted a lot of traffic and brought huge profit margins. Based on the above background, this study combines the heuristic-systematic model of information processing to classify the heuristic and systematic factors that influence the dissemination of WeChat push. Analyze the factors affecting WeChat push transmission, supplement relevant theories, and provide suggestions for WeChat Subscription Accounts operators.

1. Introduction

With the advent of the 4G, people are immersed in the massive information and convenient life brought by smart phones. The influence of mobile Internet on people's daily life is also deepening. The academic concept of "We-Media" as a new field of communication has become people's daily.

Since the media carrier is diverse, the most familiar and most frequently is WeChat Subscription Accounts. It was launched in August 2012. As of July 31st, 2017, there were more than 20 million WeChat Subscription Accounts. Through the account, enterprises can realize and communicate with the specific users of the Subscription Accounts on the WeChat public platform. This also enables some merchants to use the WeChat public account to form an official online and offline interactive marketing method.

With the rapid expansion of the WeChat Subscription Accounts, the number of followers has also risen geometrically. The WeChat subscription Accounts has also frequently produced "explosive texts". Articles on social hot issues can influence the public opinion and even affect the judicial process. The lower entry threshold and high income make more and more people work on the WeChat Subscription Accounts. In order to achieve high profit by the public number, most operators of the public number want their own WeChat Subscription Accounts has more "explosive texts". So the popularity and influence of WeChat Subscription Accounts is greater and wider, and it has a stronger ability to liquidate. However, what factors affect the effective transmission of the WeChat Subscription Accounts is and how to improve the read and share count of the WeChat push is a problem that operators and scholars have been paying attention to. So this paper studies the factors that influence the dissemination of WeChat Subscription and its inherent law.

2. Literature Review

In order to make a more systematic study on the factors affecting the dissemination of WeChat subscription Accounts, this paper studies and summarizes the literature on heuristic-systematic models, we-media, content analysis and WeChat dissemination to provide theoretical support.

2.1 Heuristic-systematic model

The basic theoretical model used in this study is the Heuristic-systematic model of information processing (HSM), which was proposed by Chaiken in the 1980s and is mainly used to explain the individual's model of information processing behavior[1]. Steginga et al. (2004)studied the utility of HSM models in hospitals for patient decision making[2]. Wirth(2007) proposed that network information search behavior can be decomposed into heuristic and systematic modes[3]. Lucassen(2011) pointed out that students with higher trust in Wikipedia tend to use heuristic information processing mode, pay more attention to the amount of information, and vice versa, prefer to use systematic information processing mode[4]. Tang and Chen used heuristic-systematic models to study the influencing factors of the dissemination of official WeChat subscription account in colleges[5].

2.2 We-Media

We-Media is a concept introduced by scholars in the analysis of blog development. It is a general term for new media that communicators with privatization, civilianization, generalization, and autonomy use modern and electronic means to conveys normative and non-normative information to specific individuals[6]. Some scholars take Weibo as an example to study the characteristics of we-media communication. They believe that the self-media communication subject represented by Weibo is more aware and can choose the audience independently. The content of the dissemination is mainly based on short fragmented information. The effect of communication can form powerful public opinion waves in an instant[7][8]. The current we-media has changed from individual to "organizational", media carriers are more diverse, and the content of communication is more diverse[9].

2.3 Content analysis

Content analysis is a research method for disseminating information content[10]. It is a method for systematic, objective and quantitative analysis of transmission for the purpose of measuring variables[11]. Liu believes that content analysis is helpful for information processing of online public opinion[12]. Wu and Rao studied the theme of urban government affairs public number based on content analysis[13]. Wang used content analysis to evaluate and study the hot content published by the library WeChat subscription account[14]. Yin and Li used content analysis to study the dissemination effect of WeChat subscription account "Tencent Public Welfare"[15]. Jia and Gu used the content analysis to study the framework construction characteristics and strategies of the government WeChat using the "Beijing Publishing Hall" as an example[16].

2.4 WeChat -IT enabled collaboration

Good collaboration is the prerequisite for completing the task in the organizations. Good information exchange is the basis of effective collaboration. In order to accomplish the task, the members of both sides will share and collaborate in the use of information resources and information technology, which makes information collaboration an important topic[17]. As we all that WeChat is an online communication and collaboration tool. It could help the development of IT enabled collaboration and communication.

Jing believes that the dissemination of the WeChat public platform has brought about information customization and private communication, opening a new network mass communication model[18]. Zhang began with the theory of complex networks to study the influence of the topology of WeChat networks on network information dissemination[19]. Wu found the impact of the keyword and semantics of the title on WeChat propagation through the analysis of the specific WeChat subscription account[20]. Bai believes that the WeChat subscription account has practical application value in the process of news dissemination[21]. Many scholars use a specific WeChat subscription account as an example to study the factors affecting the spread of WeChat. Qi and Chen believe that the WeChat public number of journals in journalism and communication should focus on operational management, content editing and structural functions[22]. Zhang used "Xinhua News Agency" as an example to propose the strategy of making WeChat push headline titles[23]. Wang conducted an empirical study on the subject of the dissemination of the article on the WeChat subscription account of the library[24]. Tian believes that the main factors affecting WeChat marketing communication in the publish are brand positioning, theme content and interactive methods[25].

2.5 Innovation

According to the literature research, we can find that the shortcomings of the existing research are mainly reflected in:

The current research on the dissemination of WeChat subscription accounts has adopted fewer theories, and there are fewer documents using heuristic-systematic model. Most of the scholars use case studies, questionnaires and interviews to establish stereotypes. The conclusions are mainly based on the personal feelings of the respondents. And the amount of data is small and representative is not strong, and it is not scalable.

The innovation of this paper is:

First, heuristic-systematic information processing models are mostly used in psychology and traditional communication studies. This paper uses this model to provide theoretical supplements for the related research on the influencing factors of WeChat push.

Second, the previous research on the WeChat subscription account is mostly qualitative analysis, or based on a specific case. This paper through the cooperation with Sootoo Netwok Company broadens the data source, using quantitative empirical analysis to study the influencing factors of WeChat push dissemination.

Thirdly, the previous research on the WeChat subscription account mostly focused on the time, form or title of the article. There are few quantitative studies on the content characteristics. This paper uses characteristic annotation, and cluster analysis, through quantitative empirical research to explore the content characteristics of the impact of the explosion.

3. 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: "entertainment", "emotion" and "news". The amount of data is shown in Table 1.

Table 1. Data overview

Category	Name	Push amount	Fans amount	New fans amount	Reading amount	Sharing amount	Likes amount	Favorites amount	Comments amount
Entertainment	Movie Heaven	547	1252926	110207	11438334	122595	87263	37377	26671
	Creative Agency	790	1202837	80647	16149409	432525	197349	78013	29044
	Daily Yoga	893	593335	167649	8244631	214584	36631	127608	21010
Emotion	Mood Signature	782	980571	62634	12568199	223379	82108	59302	27088
	Fortune Telling	318	1001620	13962	35456793	43998	13144	8342	4756
	Horoscope	594	401446	73638	5340798	48548	12245	10546	3462
News	Human Resource	838	1085823	224822	8459844	497414	23347	164530	5851
	Financial Reference	255	228286	4984	1148127	23172	5752	2362	988
	Headline	596	201849	4450	1036614	52614	5008	4159	1094

3.2 Variable definitions

Based on the heuristic-systematic model of information dissemination, combined with the characteristics of article dissemination of Wechat Public Number, this study clearly defines the impact of Wechat Public Platform push articles from two factors: heuristic factors and systematic factors. This paper divides the variables into heuristic variables, including the degree of acceptance of articles and the objective conditions of articles. The former includes the number of fans in public numbers, the rate of forwarding articles, the number of readers in friends circles, the number of readers in friends' forwarding and whether it is related to the topic of public numbers. The latter includes whether headlines, the order of pushing articles, whether weekends, and the time of pushing articles. Whether it is a popular paragraph and whether the title contains hot spots or not, systematic factors include the form and content characteristics of the article. The former includes the number of pictures and text contained in the push content, while the latter divides the content characteristics of the article into conformity with the theme, innovation, entertainment, usefulness, storytelling, value and whether to promote advertising.

The index of this paper is mainly used to measure the dissemination effect of Wechat public signal push. Broadly speaking, the dissemination effect refers to the objective results caused by the dissemination behavior, including all the impacts and results on the surrounding population and society. In narrow sense, it refers to the extent to which the disseminator's certain behavior achieves its intention or goal. This study uses a narrow definition to measure the effectiveness of disseminators' information transmission. According to the characteristics of Wechat communication, we use the amount of reading and sharing to represent the breadth of communication, and use the amount of praise and comment to represent the depth of communication.

Table 2. Variable definition details

Independent variable		Name of the variable	Meaning	Type
Heuristic factors	Push acceptance	is_tittle_theme	The title accord with the topic	0-1
		total	Instant fans amount	Integer variable
		share_rate	Push sharing rate	Fractional variable
		feed_read_count	Number of reading friends	Integer variable
		friends_read_count	Number of readers after sharing	Integer variable
	Objective condition	is_weekend	Push on weekends	0-1
		is_rush_hour	Push during peak hours	0-1
		is_first	Headline	0-1
		idx	Push order	Integer variable
		is_tittle_hotspot	Title contains hotspots	0-1
Systematic factors	Push presentation	img_num	Number of images	Integer variable
		word_count	Content words	Integer variable
	Content characteristics	is_theme	Fit with the theme	0-1
		is_creative	Innovation	0-1

Name of the variable	Meaning	Type
is_fun	Entertainment	0-1
is_useful	Usefulness	0-1
is_story	Story	0-1
is_ad	Advertisement	0-1
Name of the variable	**Meaning**	**Type**
read_count	The amount of reading	Integer variable
share_count	The amount of sharing	Integer variable
like_count	The amount of likes	Integer variable
comment_num	The amount of comments	Integer variable

3.3 Data process

3.3.1 Outlier removal

There may be abnormal data in the data collected in this paper, such as "explosive texts" 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.

Taking the entertainment subscription account as an example, the reading scatter plot of Movie Heaven, Creative Agency and Daily Yoga are shown in the figure:
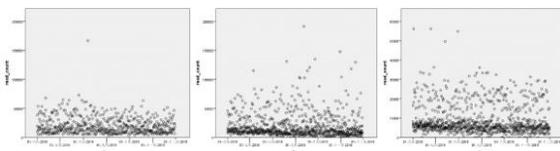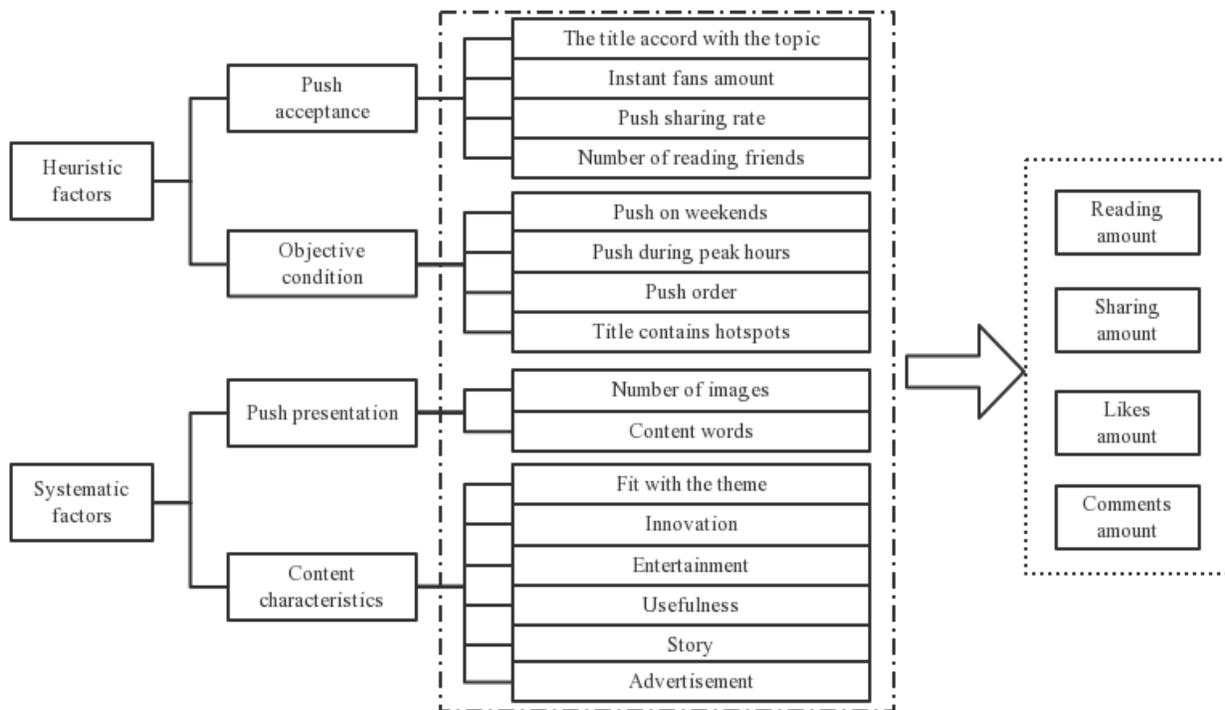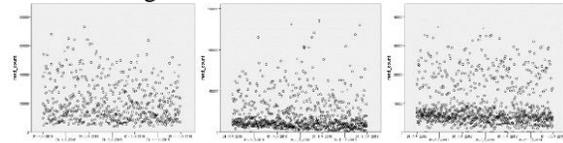
After removing the outliers:

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

3.3.2 Elimination of collinear variables

In order to avoid the problem of multi-collinearity, this section firstly calculates the Pearson correlation coefficient on the data of the WeChat push. And filters out the data items whose correlation (the absolute value of r) is greater than 0.7. According to the above correlation analysis, the updated theoretical model is shown in Figure 2.

Figure 2. Basic model

4. Data Analysis

In this chapter, the independent variables are divided into heuristic and systematic, the influence on the dissemination of WeChat subscription articles is discussed from the four dimensions, as reading volume, sharing volume, praise volume and comment volume.

To further explore the effects of different types of independent variables on dependent variables, we first make a descriptive statistical table of dependent variables:

Table 3. Dependent variable description Statistical table

	Number of cases	Minimum	Maximum	average	Standard deviation	variance
read_count	5532	0	134996	11745.75	12539.906	157249250.5
share_count	5532	0	4962	274.79	486.715	236891.432
like_num	5532	0	2946	79.89	151.556	22969.107
comment_num	5532	0	442	20.84	32.649	1065.984
Number of valid cases	5532					

From the above table, we find that the variance of the dependent variable is greater than the expected value, so we construct reading, sharing, praising and commenting negative binomial regression model in this section is as follows:

$$\ln E(read\_count) = A \times E_1 + \alpha + \varepsilon_i$$

$$\ln E(share\_count) = B \times E_2 + \alpha' + \varepsilon_i'$$

$$\ln E(like\_num) = C \times E_3 + \alpha'' + \varepsilon_i''$$

$$\ln E(comment\_num) = D \times E_4 + \alpha''' + \varepsilon_i'''$$

Among them, A, B, C, and D are the influence coefficients of the variables on the dependent variable,

E1 is the heuristic independent variable set, and E2, E3, and E4 are the heuristic and systematic variable sets, which are constant terms and are random items.

In the original data of this study, the internal discrepancies of the three independent variables, such as the amount of instant fans, the forwarding rate, and the number of readings in the circle of friends, are very large. Therefore, the values of the three independent variables are firstly processed in a logarithmic manner to reduce the impact of different orders of magnitude. Then use Negative binomial regression on WeChat subscription number. The results are as follows:

Table 4. Negative Binomial Regression Model Processing Result

Dependent variables	Model variables	Negative binomial correlation coefficient		
		entertainment	emotion	News
Reading volume	Whether the title matches theme	0.00169***	0.364***	-0.0950**
	Instant fan volume	1.457***	1.089***	1.572***
	Forwarding rate	7.466***	13.25***	7.652***
	Friend circle readings	0.377***	0.471***	0.733***
	Whether weekend	0.0105	0.0292	0.0473
	Whether to push 17-22 points	-0.176*	-0.0692**	-0.00316
	Push order	-0.296***	-0.244***	-0.312***
	Whether the title contains hotspots	0.0966***	-0.103*	-0.0985*
	_cons	0.972*	2.478***	-1.400***
	Wald chi2(8)	6672.12	3594.88	7149.9
	Prob > chi2	0	0	0
Sharing volume	Whether the title matches theme	0.0835***	0.0244	-0.149**
	Instant fan volume	0.870***	1.100***	1.259***
	Forwarding rate	27.87***	32.08***	14.74***
	Friend circle readings	0.823***	0.821***	1.075***
	Whether weekend	0.024	0.0167	0.0573*
	Whether to push 17-22 points	-0.194	-0.0500*	-0.00166
	Push order	-0.224***	-0.307***	-0.168***
	Whether the title contains hotspots	0.0473	0.00472	-0.0802*
	Number of pictures	0.116**	0.133**	-0.131**
	Article word count	0.0208	0.0411*	0.0269
	Subject fit	0.112***	0.400***	0.0125
	Innovative	0.041	0.204	-0.022
	Entertaining	0.0961***	0.0484	-0.0497
	Usefulness	0.150***	0.647***	-0.0222
	Story	-0.0422	0.0988	-0.108***
	Promoting ads	-0.232***	-0.305**	-0.423***
	_cons	-2.321***	-3.883***	-5.301***
	Wald chi2(16)	14234.59	14242.1	21277.98
	Prob > chi2	0	0	0
Like volume	Whether the title matches the theme	0.172***	-0.0395	-0.399***

	Instant fan volume	3.015***	1.856***	1.298***
	Forwarding rate	-4.776***	-2.423	-1.105
	Friend circle readings	0.524***	0.516***	0.563***
	Whether weekend	-0.0854**	0.00584	-0.103*
	Whether to push 17-22 points	-0.272	-0.0391	-0.297
	Push order	-0.381***	-0.268***	-0.430***
	Whether the title contains hotspots	0.05	0.095	0.392
	Number of pictures	0.12	0.271***	-0.502***
	Article word count	-0.226***	-0.167***	0.299***
	Subject fit	0.0449	0.511*	-0.0987
	Innovative	0.296***	0.620**	-0.0352
	Entertaining	0.142**	0.990***	0.124
	Usefulness	-0.0385	1.412***	-0.368***
	Story	0.145**	0.826***	0.0986
	Promoting ads	-0.562***	-0.555**	-0.346*
	_cons	-13.17***	-8.054***	-4.888***
	Wald chi2(16)	8902.21	7007.54	1660.73
	Prob > chi2	0	0	0
	Whether the title matches the theme	0.094	0.0619	-0.171*
	Instant fan volume	2.335***	2.951***	1.528***
	Forwarding rate	-6.032***	-0.54	-7.632***
	Friend circle readings	0.355***	0.490***	0.717***
	Whether weekend	-0.0802	-0.0672	-0.0271
	Whether to push 17-22 points	0.368***	0.114	-0.0528
	Push order	-0.425***	-0.431***	-0.446***
	Whether the title contains hotspots	0.178**	0.129	0.105
	Number of pictures	-1.367***	0.195	0.108
Comment volume	Article word count	0.0273	-0.380***	-0.359*
	Subject fit	0.163**	0.551*	0.0443
	Innovative	0.0948	0.672**	0.0343
	Entertaining	0.0356	0.770**	0.198
	Usefulness	0.459***	1.112***	-0.369***
	Story	-0.0182	0.772***	0.302**
	Promoting ads	-0.186***	-0.206***	-0.593***
	_cons	-9.264***	-14.61***	-6.771***
	Wald chi2(16)	2181.01	2469.07	1192.16
	Prob > chi2	0	0	0

Combined with the above table and analysis, we can draw the following conclusions:

4.1. Heuristic factors

(1) The number of readings in the circle of friends has a positive impact on the amount of reading, sharing, praise, and comments. The forwarding rate has a positive impact on the amount of sharing.

(2) For the entertainment WeChat subscription number, the article with the title and the theme is conducive to increasing the amount of sharing and praise of the article. For the emotional WeChat subscription number, the article with the title and the theme can improve the reading of the article.

(3) As the push order is lower, the amount of reading, sharing, likes, and comments is also significantly lower, indicating that the more forward pushes of the same WeChat subscription number on the same day are more likely to be spread more widely.

(4) The inclusion of hot content in the title has a positive effect on the dissemination of entertainment WeChat subscription articles, has a negative effect on the reading and sharing of information articles. And has little effect on the spread of emotional articles.

4.2. Systematic factor

(1) For the emotional WeChat subscription number, too many words in the article will have a negative impact on the like number and comments, indicating that too long articles may reduce the reader's behavior of sharing, likes and comments.

(2) For the entertainment WeChat subscription number, the entertainment and useful content related to the WeChat subscription concern area has a positive influence on the sharing, while the innovation, entertainment and story have positive impact, relevance and usefulness to the area have a positive impact on the review.

(3) For the emotional WeChat subscription number, the content of the article is consistent with the theme has a positive influence on sharing. The entertainment, usefulness and story have positive influence on the likes, usefulness and storytelling have a positive impact on Comments.

(4) For the information WeChat subscription number, the storyline has a positive impact on the number of comments.

(5) Advertising content has different degrees of negative impact on sharing, likes and comments.

Although the WeChat subscription number is indispensable for advertising t, it does affect the dissemination of articles.

5. Push Content Hotspot Prediction Model and Application

In this chapter, we will select the hotspot and freezing point article data, cluster the characteristics of the systemic independent variables, explore the model of the content hotspot, and use the test set to test the model.

5.1 Data description

Based on the data we have selected in Chapter 4, we forecast the content hotspots. First let's look at the description table of all the usage data, as shown in the following table.

Table 5. Readings and sharing statistics of all pushed articles

		read_count	share_count
Number of case	Effective	5532	5532
	Missing	0	0
Percentile	25	4006.75	34
	50	7815	112
	75	14649	284.75

We define articles that push hot content for both reading and forwarding. In this study, we define the articles in the latter 75%, that is, the readings are greater than 14649, the forwarding volume is greater than 284.75, 842 articles in total. We define the push freezing articles with low reading and forwarding volume, they are defined as the top 25% of the articles in this study, that is, articles with less than 4006.75 readings and less than 34 forwardings, totaling 757 articles.

Table 6. Data

Variable	Value
Total number of pushes	5532
Hot articles	842
Freezig point articles	757
Selected hot articles	600
Selected freezig point articles	600

5.2 Experimental design

In order to avoid over-fitting, this experiment uses simple cross-validation, that is, from all the 1200 selected data, according to whether the article is hot content, 400 pieces of 800 data are selected as the training set, and the remaining 200 pieces of 400 are test sets. All systemic independent variables are used as feature variables. The total feature variables are: number of pictures, number of articles, subject fit, innovation, entertainment, usefulness, story, and advertising.

During the experiment, we clustered the training sets of the two types of articles separately, and used the cluster center points as the feature center of each type of article, and then used the article feature variable data in the test set to calculate the two types of article feature centers' Euclidean distance, using the idea of the KNN algorithm to determine which type of article feature center point is close to the test article, we think that the test set article belongs to which, we record it and compare it with the known results to judge the effect of the cluster prediction model.

5.3 Experimental process and results

In the experiment, we regard each article as a point in the N-dimensional space whose coordinates are the space vectors formed by the feature set variables. Firstly, the feature variables of the training set data are standardized to meet the requirements of K-means clustering. Then, K-means clustering is performed in the training set of hot articles, so that the K value is tested from 2 to multiple times, until the clustering result can make as many cases as possible, taking the clustering result as the model result, and regard the center point of the most individual as the content feature center point of the hot article, clustering of the hot article. The result is as follows:

Table 7. Hot Article Training Set's Final Cluster Center

	Cluster		
	1	2	3
Zscore(img_num)	0.0114	-0.71406	-0.05801
Zscore(word_count)	-0.06387	0.72301	6.87787
Zscore(is_theme)	-0.00079	-0.52658	1.15578
Zscore(is_creative)	0.0077	-0.58816	0.17321
Zscore(is_fun)	0.01398	-0.6074	-0.6074
Zscore(is_useful)	0.0025	0.04662	-0.41956
Zscore(is_story)	0.00448	-0.31423	0.04489
Zscore(is_ad)	-0.12325	8.09336	-0.12325

Table 8. The number of individual cases of hot articles in each cluster

Cluster	1	391
	2	6
	3	3
Effective		400
Missing		0

In the course of the experiment, when k=3, we found that there is no change in the cluster center after 4 iterations, and a good convergence is achieved. The clustering results at this time have also reached our goal that is having 97.75% of the number of cases in a class. We believe that the center of this class (cluster 1) can be regarded as the center of the feature variable of the hot article. Similarly, we can also make clustering

results for freezing articles, as shown in the following table:

Table 9. Freezing point article training set's final cluster center

	Cluster	
	1	2
Zscore(img_num)	0.09376	-0.5881
Zscore(word_count)	-0.20987	1.31648
Zscore(is_theme)	-0.05702	0.35766
Zscore(is_creative)	0.02804	-0.17587
Zscore(is_fun)	-0.00879	0.05513
Zscore(is_useful)	0.07498	-0.47035
Zscore(is_story)	-0.32359	2.02979
Zscore(is_ad)	0.14185	-0.8898

Table 10. The number of individual cases of freezing point articles in each cluster

Cluster	1	345
	2	55
Effective		400
Missing		0

In the experiment, we found that there is no change in the cluster center after 8 iterations at k=2, and a good convergence is achieved, and there are 86.25% of the cases in one class, and there is no increase in the value of k. So we think that the center of cluster 1 can be regarded as the center of the feature variable of the freezing point article.

From the above steps, we have calculated the center of the hot article and the freezing point article. We can calculate the Euclidean distance between the normalized 400 test individuals and the two center points. By comparing the two distances, the test individual is divided into the class of the center point with a small distance, and compared with the actual situation, the confusion matrix of the test result is shown in the following table:

Table 11. Confusion matrix of hotspot prediction results

	Actually hot	Actually freezing point
Forecast hot article	158(TP)	56(FP)
Forecast freezing point article	42(FN)	144(TN)

5.4 Experimental evaluation index analysis and model application

In order to evaluate the evaluation effect of the clustering model, this section selects the correct rate, accuracy, recall rate, and F1 score to evaluate the model. The calculation methods of these evaluation indicators are related to the four statistics of the confusion matrix, combined with the experiment. In terms of design, the real class (TP) is the actual hot article and is predicted as the sample size of the hot article. The false positive class (FP) is the sample size that is actually the freezing point article and is predicted to be a hot article, and the fake negative class (FN) It is a sample size that is actually a hot article but is predicted to be a freezing point article. The true negative class (TN) is a sample size that is actually a freezing point article and is predicted to be a freezing point article.

The evaluation indicator formula used is as follows:

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

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

Recall：
$$\mathrm{Re} call = TP/(TP+FN) \times 100\%$$

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

From the above formula combined with the research variables of this experiment, the correct rate measures the percentage of the total number of samples that are correctly predicted and classified in all individuals. The accuracy rate indicates that all the clustered models are predicted to be hot articles, also for the proportion of hot articles. The recall rate measures the proportion of samples that are predicted to be hot articles for all actual hot articles. It is known from the statistical law that the accuracy rate and the recall rate have a trade-off relationship, and the F1 score is a comprehensive indicator that considers both the accuracy rate and the recall rate. The evaluation indicators for this experiment are shown in the following table:

Table 12. Experimental evaluation index

index	Accuracy	Accuracy	Precision	F1 score
value	75.50%	73.83%	79.00%	76.33%

As can be seen from the above table, all the evaluation indicators of the model are above 70%, indicating that the performance of the prediction model is relatively balanced, and the prediction effect is good and acceptable.

The design of this model is to help the WeChat subscription operator to better identify whether the article meets the characteristics of the hot article, and to avoid the freezing point article, so that the WeChat subscription article can spread better and enhance the value of the WeChat subscription account.

6. Summary and outlook

Based on the real operational data of 9 WeChat subscription accounts, this paper supplements the blank of WeChat research by using the heuristic-systematic model of information processing, and draws the following conclusions:

6.1. Heuristic factors

(1) In terms of the title of the article, the title of the push is in line with the theme of the WeChat subscription account, and the inclusion of characteristic words has a certain positive influence on the dissemination of the article. For the entertainment and emotional WeChat subscription numbers, the title is compatible with the field of interest of the WeChat subscription number, and the article with the hot topic is better. For the information WeChat subscription number, only the praise and the title are related to the WeChat subscription. The relevant field, and the article containing hotspots is not conducive to the dissemination. This may be different from the user's attention characteristics. Users who are concerned about entertainment and emotions are more inclined to entertaining, so they want to read and share articles related to the topic. For the information WeChat subscription number, the followers may more inclined to understand the information content, rather than just relying on the title to determine whether to read and share and other user information dissemination behavior.

(2) The high forwarding rate and the high circle of reading have a positive impact on the dissemination of the article. With the increase of the forwarding rate and the reading volume of the circle of friends, there is a positive correlation between the reading and sharing of the three types of WeChat subscription accounts. Both of these independent variables can explain the reader's attitude. It also measures the breadth of the dissemination of articles, indicating that articles approved by friends will lead to better communication effects, thus forming a virtuous circle, making the WeChat subscription number more and more widely spread.

(3) In terms of objective conditions of the article, the headline article has obvious advantages. In the entertainment WeChat subscription account, the reading and sharing of headlines and sub-articles are significantly higher than those of Articles 3-6, while in the emotional and information WeChat subscription accounts, only the headline readings and sharing amounts are significantly higher than other pushes, this shows that the entertainment article WeChat subscription number is also very attractive to readers.

6.2. Systematic factor

(1) In terms of the form, the articles with more pictures are obviously more shared and praised than the articles with fewer pictures, while the articles with moderate number of words have more praise and sharing. It shows that multi-graphs are more likely to allow users to disseminate than multi-words, while lengthy articles are mostly boring, and readers are less willing to share.

(2) In terms of the content characteristics, the influence of different article content characteristics on different types of WeChat subscription numbers is not the same. In the entertainment WeChat subscription account, the theme, entertainment and usefulness are important; in the emotional WeChat subscription account, story, usefulness and entertainment are more popular among readers. The articles have more forwarding numbers, likes and comments; in the information WeChat subscription number, the articles will be more shared and commented that they match the theme of the WeChat subscription, usefulness, and story.

(3) The advertising content has a strong negative impact on the communication effect of the three types. It is understandable that ordinary users do not like to receive advertising content and are not willing to actively distribute advertising content.

In addition, this paper also gives a general method for the establishment of predictive models based on the existing research data and the freezing point article, and analyzes the effect of the forecasting model. In the future, the operator of the WeChat subscription account can establish a certain forecast hotspots and freezing point models that belongs to the unique account to help them use more scientific methods and more comprehensive indicators to identify potential push heat.

7. Acknowledgment

References

[1]Chaiken S. Heuristic Versus Systematic Information Processing and the Use of Source Versus Message Cues in Persuasion [J]. Journal of Personality & Social Psychology, 1980, 39(5): 752-766.

[2]Steginga, Suzanne K.; Occhipinti, Stefano. The Application of the Heuristic-Systematic Processing Model to Treatment Decision Making about Prostate Cancer [J]. Med Decis Making. 2004, 24 (6): 573–583

[3]Wirch W，Becking T，Karnowski V，et al．Heuristic and systematic use of search engines [J].Computer-Me-diated Communication, 2007(3):778-800

[4]Lucassen T, Noordzij M L, Schraagen J M. Reference blindness: The influence of references on trust in Wikipedia[C] //Proceedings of the ACM Web Science

Conference Germany, 2011:1-4

[5]Tang Yayang, Chen Sanying. Empirical Study on the Influencing Factors of the Communication Effect of Official WeChat subscription Signals in Universities and Colleges [] J. Journal of Hunan University (Social Sciences Edition), 2018 (5): 155-160

[6]Dan G. We the Media[M]. O'Reilly Media, Inc. 2004.

[7]Zhang Meiling, Luo Yi. Analysis of the characteristics and advantages of self-Media communication represented by micro-blog [J]. Journal of Hubei Vocational and Technical College, 2011, 14 (1): 45-49

[8]Li Yongzhi. Exploration of the characteristics and advantages of self-Media communication represented by Weibo [J]. New Media Research, 2016,2(5): 11-11,49.

[9]Kelly Geng. The Concept and New Characteristics of Self-Media [J]. New Media Research, 2017, 3 (14): 5-6.

[10]Wu Shizhong. Outline of Content Analysis Method [J]. Information and Data Work, 1991, (2): 37-39.

[11]Li Benqian. Describing the characteristics of communication content to test the hypothesis-content analysis method of communication research (Part I) [J]. Contemporary Communication, 1999, (6): 39-41.

[12]Liu Yi. The application of content analysis in the analysis of network public opinion information [J]. Journal of Tianjin University (Social Science Edition), 2006, 8 (4): 307-310.

[13]Wu Chaoyan, Rao Yanghong. Research on the topic of information dissemination of WeChat subscription Signal in urban government based on content analysis [J]. Modern Information, 2017, 37 (2): 52-56, 67.

[14]Wang Kang, Wang Xiaohui. The status quo and Countermeasures of WeChat subscription Number Service in University Libraries [J]. Digital Library Forum, 2017, (6): 53-57.

[15]Yin Zhangchi, Li Baixuan. Study on the Communication Effect of WeChat subscription Signal -- "Tencent Public Welfare" as an example [J]. New Media Research, 2018, 4 (11): 6-8, 35.

[16]Jia Zhemin, Gu Xiaoyu. The Framework Construction and Impact of Government WeChat Communication [J]. Journal of Beijing University of Aeronautics and Astronautics (Social Science Edition), 2018, 31 (1): 32-38.

[17]REDDY J C，JANSEN B J. A model of understanding collaborative information behavior in context: a study of two healthcare teams ［J］. Information Processing and Management，2008，44 (1) : 256-273.

[18]Jingming, Zhou Yan, Ma Danchen. WeChat communication mode, characteristics and reflection [J]. Journalism and Writing, 2014, (7): 41-45.

[19]Zhang Hong, Li Jie. Research on Information Dissemination in the Complex Network of Wechat [J]. Science and Technology and Engineering, 2014, 14 (17): 244-247.

[20]Wu Zhongtang, Liu Jianhui, Tang Zhenhua. Research on Influencing Factors of Information Dissemination of WeChat subscription Number [J]. Intelligence Journal, 2015, (4): 122-126.

[21]Bai Mingrui. Analysis of the Application Value of WeChat subscription Signal in News Communication [J]. New Media Research, 2017,3(3): 1-2.

[22]Geng Rui, Chen Qian. Reflections on the Hot Construction of WeChat subscription Platform for Journalism and Communication Academic Periodicals [J]. Research on Chinese Science and Technology Periodicals, 2017,28(1): 53-57.

[23]Zhang Lingling. Research on the Making Strategy of WeChat subscription Number Title --- Take Xinhua WeChat subscription Number as an example [J]. News Knowledge, 2018, (9): 60-62.

[24]Wang Lei. Research on Twitter Theme Dissemination on WeChat subscription Platform [J]. Modern Information, 2018, 38 (10): 132-141.

[25]Tian Jun. How the Publishing Group Does WeChat Marketing Communication Well [J]. Young Journalist, 2018, (26): 62-63.