Personal and Ubiquitous Computing
https://doi.org/10.1007/s00779-023-01710-7

ORIGINAL PAPER

Exploring information dissemination effect on social media: an empirical investigation

Bo Yang[1] · Rong Zhang[1] · Xusen Cheng[1] · Chuang Zhao[1]

Received: 26 October 2022 / Accepted: 29 January 2023

Abstract
Operators make profits by publishing information. If they can know the influencing factors in the process of information dissemination, they can provide new insights for practical operations and formulate corresponding operation strategies for different types of accounts. The purpose of this article is to discuss the information dissemination process of WeChat public accounts and what factors will affect the reading rate and sharing rate of the article. In this paper, the "feedback-sympathize-identify participant-share" (FSIPS) two-stage model is used to analyze the characteristics of information dissemination, and the negative binomial regression model is used to analyze which factors have a significant impact on the two stages of the dissemination model. Our data is obtained from a company that specializes in operating WeChat Official Accounts, and the data is authentic and more valuable. We collectively consider the roles of users, message content, interactions between content and individuals, and the context of social media in information dissemination behaviors (i.e., reading and sharing). In this process, some new variables, such as environment-related variables, are involved and incorporated into the two stages of information dissemination for analysis. The results show that the like rate, which is one of the feedback dimensions, has the greatest impact on the reading rate, while the favorite rate has the greatest impact on the sharing rate. Article type also plays a crucial role in the dissemination of information. In addition, the content relevance between the title and the content also largely affects the share rates of the three types.

Keywords FSIPS two-stage model · WeChat Official Accounts · Information dissemination

1 Introduction

Social media has been embedded in people's lives through the construction of a new social structure dependent on innovative information technology [1]. It has also enabled online sharing to become a daily activity [2]. Large-scale use promotes information dissemination, which may produce information dissemination effects, that is, the impact and results of information dissemination or information dissemination process. Large-scale dissemination resulting from adoption will lead to the phenomenon of virality [3] or form a social epidemic [4]. Such phenomena have become commercial objectives affecting the formulation and implementation of marketing strategies, and they benefit the market. Widespread dissemination of information is equal to the increased number of visitors; thus, publishers can more easily acquire potential subscribers or consumers [5]. The commercial value will also be produced from triggering widespread information (e.g., social epidemic, virality), such as helping enterprises or retailers promote their brands, and obtaining potential customers more easily with lower cost [4, 6, 7].

Users or operators write posts or articles, and friends or strangers subscribe to their posts to be notified about recent updates [8]. Reading and sharing rates are important indicators of the success of a piece of content [3, 7, 9], which is supported by the follower base. Although companies have invested more resources in creating interactive content to attract followers and provided financial incentives to increase their followers' base, many still find it challenging to maintain their followers' base [10].

As the main marketing channel in Wechat, WeChat Official Accounts (WOAs) face the same issue. WOAs attract users to subscribe by posting articles [5]. The usability and functionality of WOAs make it easy for users to receive information from others and post content that can be shared

✉ Xusen Cheng
xusen.cheng@ruc.edu.cn

[1] School of Information, Renmin University of China, Beijing, China

almost instantaneously to a wide audience, making it an attractive platform to content creators and media companies. However, the operation and management of the WOAs still face more significant challenges, such as the difficulty of traffic acquisition, the low viscosity of followers, the increase of homogenization content, and the rather slower reading volume. The development of the online community benefits from users' continuous participation, and economic benefits may be generated from the interaction between users and social media [11]. Thus, a comprehensive understanding of information dissemination law, including content, users, and environment dimensions, helps grasp the whole picture of the challenges already mentioned. Determining the factors leading to the spread of personal information can help managers deal with existing problems, such as developing a reasonable network marketing strategy, effectively intervening in network rumors in an emergency [4], or controlling and disseminating communications on social media [7]. Therefore, this study attempts to provide more strategic solutions as alternatives to contemporary communication methods [12].

Researchers have long been interested in studying the information dissemination for innovation adoption [13], marketing strategy [4, 14], and emergency management [15, 16]. The information dissemination behaviors, including retweeting [4, 15, 17], forwarding [8, 18, 19], and following and unfollowing [20] in the context of Twitter, Facebook, and Microblogging, have gained great attention.

Many researchers examined the driving force behind information dissemination. One research stream focus on content, which is key to users' dissemination behavior. For instance, content characteristics (i.e., anger, sadness) have an impact on sharing, and the positive content is more viral [21]. Another research stream considers users' role. Some researchers studied individual characteristics that affect message sharing, specifically with regard to who is more likely to rebroadcast the message [22]. The combined role of content and users has also been explored. Rebroadcasting decisions depend not only on message content but also on the fitting between messages and the user [7]. However, these studies are all conducted from a limited perspective. Beyond identifying characteristics of individuals and content, it is unknown whether other features or the fusion of all features will affect users' decisions to disseminate messages. In the complex operating environment, it remains unclear whether other factors should be considered regarding the challenges faced by organizers and operators. The theory-driven investigation into the information dissemination process on social media from an overall and comprehensive perspective remains far from adequate.

This study addresses the research gaps in information dissemination by constructing an appropriate framework and determining essential factors to understand the dissemination process better. Accordingly, three research questions are examined in this study:

RQ1. What are the characteristics of individual information dissemination behavior on social media?

RQ2. What factors affect the information dissemination process and which are the most influential ones?

RQ3. How do these factors affect the information dissemination process?

To answer these research questions, we focus on describing the process of users' information dissemination behavior (i.e., reading and sharing) and the key factors of information dissemination effect. We achieve that by constructing a two-stage model named FSIPS, which comprises "feedback-sympathize-identify" (FSI) and "participate-share and spread" (PS) stages. Then, we analyze the influencing factors of reading and sharing rates from two stages to provide implications for operators and managers.

In the first stage, "sympathize" affects users to click on the article page and read the content. Specifically, the user initially determines and "identifies" whether the article pushed is consistent with personal interests or meets their needs for information search. Thus, the reading rate reflects the information dissemination effect in the stage.

We consider that four dimensions—feedback, new user base (users who have recently followed WeChat Official Accounts), content perception, and environment—may affect the information dissemination in the first stage. Feedback is one of the new and important factors in information dissemination, because users may receive articles shared by others. In addition, the quality and dissemination of articles pushed by the WeChat public platform the day before will also have an impact on today's reading rate [23]. Users, especially new users, are closely related to the feedback and interaction [20, 24] that are typical user participation behaviors and play essential roles in the information spreading process [24, 25]. Thus, the new user base of WOA influences the information dissemination in this stage. Content characteristics also significantly affect users' participation behaviors [14, 21, 26, 27].

In the early stage of the information dissemination, as users acquire content perception of the article through the WOA name corresponding to article type and title, we identify article type and articles title as factors influencing the reading rate; scholars [28] have found that when the title of news articles is more than 15 characters, the reading rate is significantly improved, so we take 15 as the critical value. From a contextual perspective, as shown in Fig. 1 previous research has shown that when an account sends multiple articles at the same time, the position of the article differs in read rate, and the higher the position of the article in the list, the higher the probability of being selected [29, 30]; therefore, we consider push position. Additionally, since the frequency of users using social media platforms is different over time, we also select push time as another environment-related factor into account.

Fig. 1 Subscribing to WeChat official account articles

In the second stage, after the process of resonance and identification, users are faced with the choice of deeper social interaction with the information, including feedback and forwarding. Feedback refers to behaviors such as comments, likes, and favorites on articles. We choose the sharing rate to measure the effectiveness of information dissemination, because sharing is the last step in which users participate in information dissemination [3, 7, 9]. Feedback and content judgment may affect this stage of information dissemination. Feedback may also have an effect on users' sharing behavior; we regard it as one of the critical factors in this stage. Besides, we introduce content-related factors (content judgment) to explore whether information dissemination behavior (i.e., sharing) will be triggered by the internal content of the article after users acquire the external content (article type and article title) in the first stage.

Two dimensions of content judgment are considered in this study. First, researchers have proved that the number of multimedia and content vividness [16, 31] has significant effects on information retweeting; thus, content vividness is considered one of the key dimensions. Second, in the context of WOAs, publishers often elaborate catchy titles and eye-grabbing descriptions to increase visitor traffic [5], which may only induce users to click, but not attract users' deep interaction such as sharing or spreading content to others if users find that the content is not as interesting as the title. Therefore, we also regard content relevance (i.e., the matching degree between title and content) as a dimension of content judgment.

This study makes several important contributions to research and practice. First, to the best of our knowledge, we are one of the first to propose the FSIPS model and apply it to the information system field to systematically examine information dissemination factors. Drawing on the sympathize, identify, participation, sharing, and spread (SIPS) model, this study integrates the feedback effect that is crucial in the social media environment to explain the information dissemination process from a more comprehensive perspective. Second, we jointly consider the role of users, message content, interplay between content and individuals, and context of social media in information dissemination behavior (i.e., reading and sharing). In this process, we also considered some new variables, such as environment-related variables, and integrated them into the two stages of information dissemination for analysis. We find that content-related factors (i.e., content relevance), environment-related factors (i.e., push position), and feedback-related factors are the most influential factors. The overall architecture is shown in Fig 2.

This paper is organized as follows. First, we review the related literature in Section 2. In Section 3, we propose our model in the context of WOAs. Next, Sections 4 and 5 elaborate on the research design and the results of data analysis. Following this, the findings, as well as the theoretical and practical implications, are discussed.

2 Literature review and theoretical background

2.1 Information dissemination in social media

Dissemination, also known as diffusion, is a communication process through which a new idea, behavior, and/or technology can be transmitted from a person, an organization, or any adoption unit to another person, an organization, or any adoption unit in the social system over time through certain channels [12]. The existing related literature on information dissemination in social media can be classified into two streams as follows.

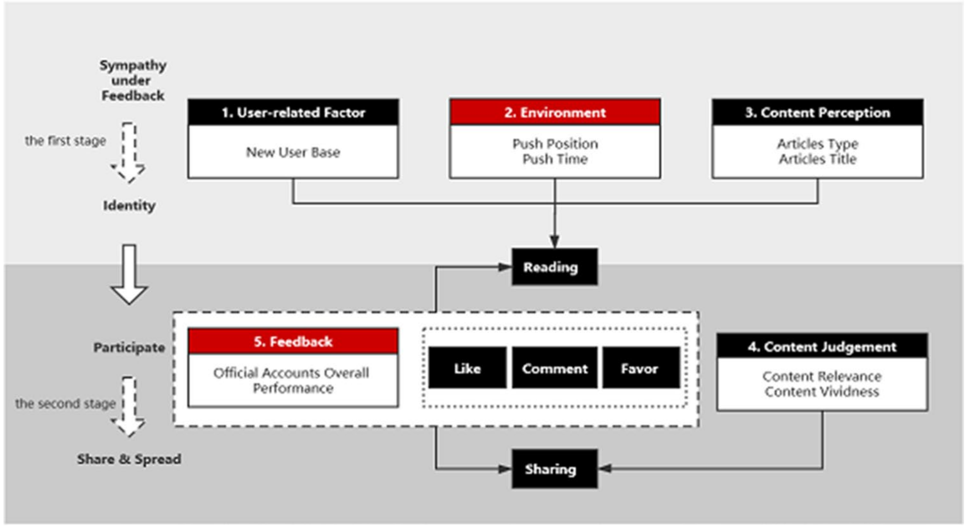

Fig. 2 Modeling framework

The first research stream emphasizes users' role in the dissemination process. User participation makes social media an important channel of information dissemination, and so users are at the core and become the entry point of research on information dissemination in social media. Stakeholder engagement is the most focused theme in this field [32]. Engagement refers to active participation in these unique sociotechnical environments based on social media platforms [33]. Additionally, users' subscription behavior is critical to information dissemination in social media. For instance, subscription behavior builds subscriber networks; thus, it plays an essential role in the early stages of information dissemination [25]. The number of subscribers or followers of an account largely determines the magnitude of information dissemination [24]. Information dissemination is heterogeneous by user type. Certain types of dominant users have been identified as information sources and information diffusers [15, 34].

The second research stream examines content characteristics and information dissemination in social media. Sharing is a typical users' behavior of information dissemination in social media. Researchers have found that content emotion [21, 26], visibility of message [14], and information value [27] significantly affect users' sharing behavior. Furthermore, there are interacting effects between creator-related factors and content characteristics. For example, Han [3] has proved the predictive validity when considering content–creator interactions in the study of virality on social media.

Although these above-mentioned studies have contributed to our understanding of information dissemination in social media, most of them have only focused on limited factors of users and content while ignoring the effects made by social media environment-related factors. Hence, it is challenging to comprehensively understand information dissemination. This research proposes a model of disseminating behavior in social media to describe the information dissemination process from a more holistic aspect and discover more comprehensive factors affecting the process.

2.2 Dissemination model

In the traditional media marketing field, the classic linear attention, interest, desire, memory, and action (AIDMA) model proposed by Samuel Roland Hall in 1924 has played a central role in describing the psychological processes and offline consumers' purchase behavior [23, 35, 36]. In response to technological changes, Dentsu[1] proposed a model containing five processes: attention, interest, search, action, and share (AISAS) based on the traditional model [37]. The AISAS model is regarded as an innovative framework for marketing communication plans because of its search and sharing mechanisms [38]. It describes online consumer behavior in the context of interactive and personalized Internet [39]. The five stages are non-linear and limitless because some information that receives online "likes" will be searched for and shared again [23], leading to information dissemination.

Many scholars have focused on the AISAS model and used it to examine online consumer behavior in social media. For instance, Tseng and Wei [37] explored the impact of media richness on consumer behavior at different AISAS stages. They found that media richness has a greater influence on the three early stages of AIS (attention, interest, search) while having a lower impact on the later stages of AS (action, share). Wei and Lu [39] compared the influence of celebrity endorsements to online customer reviews on female shopping behavior based on the AISAS model.

[1] Dentsu is the largest advertising agency in Japan.

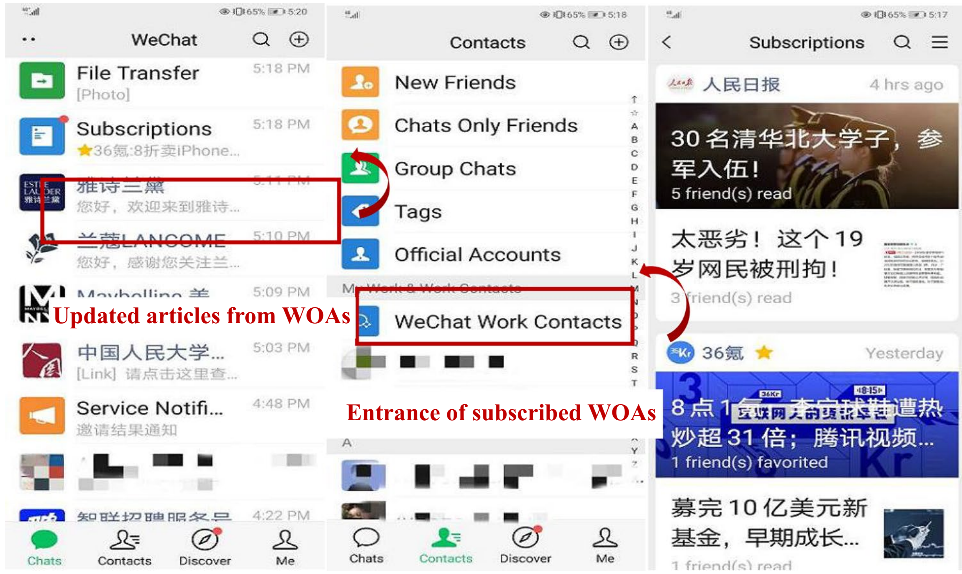

Fig. 3 Screenshot of WOAs

The AISAS model lacks interaction with customers' recognition and emotional processes [40]. With the mobile terminals bringing new changes to the information reception and dissemination of social media, Dentsu comprehensively considered user behavior in social media based on the AISAS model. Then, Dentsu proposed the SIPS model incorporating psychological and emotional factors, making it more comprehensive in describing social media user behavior.

There have been few systematic attempts to explore the information dissemination process based on the SIPS model. The SIPS model is divided into four stages: the first stage is resonance, and only when product information resonates with consumers can it further communicate and interact with the enterprise; the second stage is confirmation, where consumers resonate with themselves through external confirmation, whether the product information is valuable and eliminate the user's distrust of the product; the third stage is participation, consumers are very likely to purchase through a series of participation actions; and the fourth stage is sharing and diffusion, a good consumption experience prompts consumers to spontaneously share socially and generate secondary promotion. Although the SIPS model has been proposed in the context of online product marketing, content is also a kind of product that needs marketing strategies in social media information dissemination. Therefore, the SIPS model can help understand users' information dissemination behavior in social media and serve as the research foundation of this paper. With the development of social media, especially mobile social media (e.g., WOAs), some changes have appeared in information dissemination on social media platforms. Therefore, this study considers the characteristics of emerging mobile social media and improves the SIPS model to make it more suitable for these new characteristics.

3 Modeling framework

WOA is a function launched by the WeChat platform to convey information to users, providing a new way for media and individuals to spread information. Fig. 3 shows a set of screenshots of WOAs. Stakeholders of information circulating on WOAs include users and accounts. The latter are the source of information dissemination.

Drawing on the SIPS model, we identify two stages of information dissemination in WOAs. In the primary stage, an article is published, and followers will decide to click and read if it is interesting. The process corresponds to the two early stages of the SIPS model—sympathize and identify. Therefore, users may portray several behaviors to interact with the account or other people. That is, they may click the like button, comment, collect content to favorites, and share content to others. This corresponds to the two later stages of the SIPS model—participate-share and spread. Fig. 4 provides an overview of the process of information dissemination in WOAs.

The information dissemination process is non-linear and infinite, as new users may receive previous articles. In other words, the feedback from the user dissemination stage may influence the information dissemination behavior of new users(users who have recently followed WeChat Official Accounts).Thus, we incorporate "feedback" as a new module into the SIPS model, and construct a two-stage framework in this study (as shown in Fig. 2). Because WOAs provide users with various ways to participate and interact with information, including favorites, likes, and comments, we regard these interactions as feedback behavior. Regarding our new FSIPS model, the dissemination of articles on the WOAs is characterized by two stages of "feedback-sympathize-identify" (FSI) and "participate-share and spread" (PS).

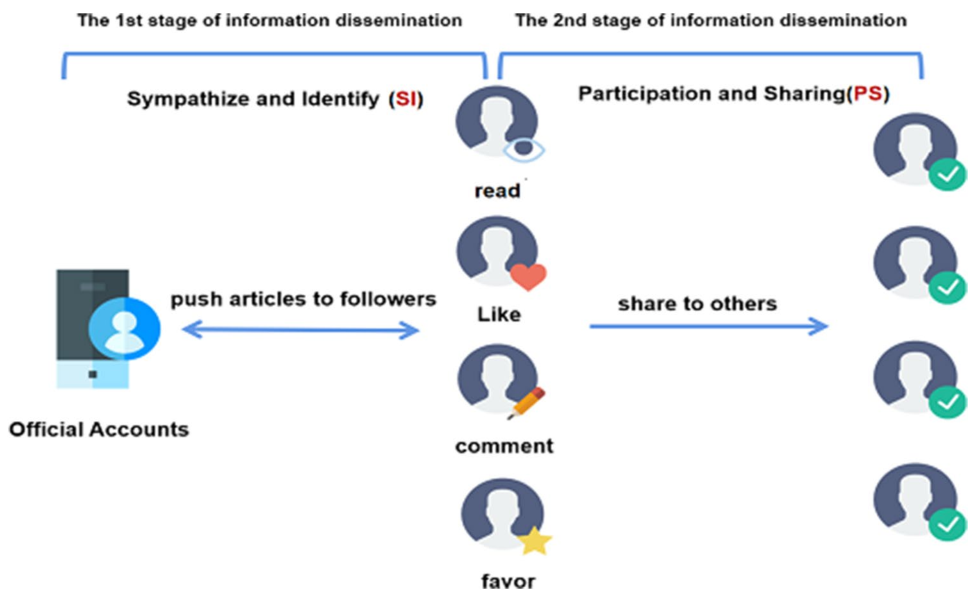

Fig. 4 Process of information diffusion in WOAs

4 Research method

4.1 Sample and data collection

With the help of Beijing Sootoo company, a brand ecological service provider based on new media and new technology, we collected data of 21 WOAs from November 30, 2017, to September 1, 2018. While 3286 articles were collected, we were left with 2969 after removing outliers (for example, the content of the article is empty, or the number of likes is greater than the number of reads) and duplicate articles. Table 1 shows 2969 descriptive statistics of 21 WOAs in three months. They contain three types of WOAs—1169 are about life, 648 are about entertainment, and 1152 are about information. We want to know the degree of influence of all the influencing factors on the information dissemination of different types of WOAs. These three types will be analyzed separately.

After the initial data processing, we use the Jieba method for word segmentation. The Jieba word segmentation algorithm uses the prefix dictionary in the data structure to scan and classify Chinese characters with high efficiency and precision. It has three modes: search engine, precise, and full modes. The first two modes are suitable for search engine segmentation and text analysis, respectively. Although the full mode is fast, it is hard to solve the ambiguity. For text analysis, the study selected the precise mode that can more accurately analyze and segment the text. Term frequency-inverse document frequency (TF-IDF) algorithm can be used to evaluate the importance of a word to a file set or one of the files in a corpus. Therefore, we can calculate the TF-IDF weights of different word segments and extract the article content keywords whose TF-IDF weights are located in the top ten. We performed negative binomial regression analysis using Stata.

WeChat users are interested in the official account. When some public operators use hyperbolic titles or hot topics to attract WeChat users, but the relevance and matching degree of the actual content and the topic are not high, users will perceive them as "unworthy of the name." This will greatly weaken their willingness to take in-depth actions (like, comment, collect, and share). Based on the 10 keywords extracted by the above algorithm, we first use Excel function to judge whether each keyword is included in the title. We then calculate the coverage of 10 keywords (keyword coverage = number of keywords included in the article title/10)×100%), and then get the matching degree between the text content and article title.

Table 1 Descriptive statistical results of the data

Variable	Obs	Mean	Std. dev.	Min	Max
account_type	2969	1.994	.884	1	3
is_includevideo	2969	.135	.342	0	1
ratio_commment	2969	.002	.003	0	.079
ratio_like	2969	.012	.015	0	.127
ratio_favor	2969	.008	.011	0	.176
read_num	2969	8237.715	10171.36	19	218000
share_num	2969	265.551	621.426	0	16091
conlen_img	2969	191.719	278.85	0	5652.5
middle_relevance	2969	.089	.285	0	1
high_ relevance	2969	.378	.485	0	1

Table 2 Dependent and exposed variable definitions

	Stage	Variables	Variables identify
Dependent variable	The first stage	Read number of the day	read_num(i,j,t)
	The second stage	Share number of the day	share_num(i,j,t)
Exposure variable	The first stage	WOAs' cumulate user of the day	cumulate_user(i, t)
	The second stage	WOAs' read number of the day	read_num(i,j,t)

4.2 Data variable definition

In the first stage, "feedback-sympathize-identify," the resonance stage is affected by new users, push environment, external content perception, and feedback. In the content confirmation stage, the article's click reading rate can reflect user participation. Therefore, the first phase of the research model is to push the click reading rate of the article as the research object. The reading number is set as the dependent variable. On the same day, the official account is concerned about the number of exposures. The potential factors are selected as the independent variables in the four dimensions of the new user base, the push environment, the external content perception, and the feedback effect. More focus is put on the behavior transformation mechanism of the "attention click reading" of WeChat users.

In the second stage of "participation-share and spread," the "participation" takes the user's click to read as the starting point and further transforms into deep participation behavior, including favor, likes, and comments. For "share and spread," the communication effect can directly and effectively be measured by the amount of sharing. A user's deep participation behavior is an important basis for sharing and diffusing information with other users. Therefore, in the second stage, the study considers the sharing amount of push articles as the dependent variable and regards the number of clicks to read as exposure variables. It then regards the favor, comment, share rates, and the intrinsic characteristics of the push content (content vividness and internal correlation) as the independent variables.

The specific definitions of dependent variables and exposure variables are shown in Table 2, and independent variables are shown in Table 3.

Table 3 Independent variable definitions

Stage	Dimension	Independent variables	Variables identify
The first stage	New user base	The proportion of new users on the day	ratio_newuser (i, t)
		The proportion of new users on the previous day	ratio_newuser_t1(i, t-1)
	Environment	Push time, whether weekend	is_weekend (i, j, t)
		Whether lunch break	is_noonbreak (i, j, t)
		Whether time from work	is_quitting (i, j, t)
		Whether before bedtime	is_bedtime (i, j, t)
		Push position whether headline	is_first (i, j, t)
	Content perception	WOAs type	account_type (i)
		Whether title length not more than 15 words	is_notmorethan15 (i, j, t)
		Whether title contains emotional punctuation	is_emotional (i, j, t)
	Feedback	Comment rate of the previous day	ratio_comment_t1 (i, t-1)
		Like rate of the previous day	ratio_like_t1 (i, t-1)
		Favor rate of the previous day	ratio_favor_t1 (i, t-1)
		Sharing rate of the previous day	ratio_share_t1 (i, t-1)
The second stage	Action participation	Comment rate of the day	ratio_comment (i, j, t)
		Like rate of the day	ratio_like (i, j, t)
		Favor rate of the day	ratio_favor (i, j, t)
	Content judgment	WOAs' type	account_type (i)
		Picture text ratio	conlen_img (i, j, t)
		Whether contains video	is_includevideo (i, j, t)
		Content middle matches title	middle_relevance (i, j, t)
		Content high matches title	high_ relevance (i, j, t)

4.3 Analysis method

In this study, Poisson regression and negative binomial regression model were used for data analysis. The basic premise of the Poisson regression model is that e (y) = var (y) =λ, which means the variance must be equal to the mean; that is, there is no excessive dispersion, and it should be "equal dispersion" [14]. However, the variance of many count data is greater than the mean value in practical, and "excessive dispersion" occurs occasionally. If we continue to use the Poisson regression model, although the consistency and unbiasedness of regression coefficient estimation will not be affected, the underestimated standard error will reduce the accuracy of the statistical test. To avoid the above problems, we consider the overdispersion effect and introduce a random error term into the original Poisson regression model, extending the model to build a "negative binomial regression model" [9]. This study applied the negative binomial regression model to adjust the analysis for overdispersion.

The main factors affecting the reading and sharing rates can be studied using a negative binomial regression. When using the negative binomial regression model for statistical analysis, we can use the alpha test to determine whether there is excessive dispersion of the experimental data. If the result of the alpha test is significant, the experimental data need to adopt a negative binomial regression model for statistical analysis.

Since the negative binomial regression is improved based on the Poisson distribution regression model, the traditional Poisson distribution is expressed as follows:

$$P\left(y_i\right) = \frac{\lambda_i^{y_i} \exp\left(-\lambda_i\right)}{y_i!}$$

$$\lambda_i = V_i$$

where $p\ (Y_i)$ represents the probability, λ_i denotes expectation, and V_i denotes variance, expressed by propagation influencing factors. λ_i formula is expressed as follows:

$$\ln \lambda_i = \ln E_i + \ln X_i \beta_i$$

where E_i and X_i represent factors affecting communication and β_i is the undetermined coefficient.

Poisson distribution requires equal mean and variance, but the actual data cannot meet this requirement; hence, the error term is added ε_i. After the adding error term, λ_i function expression and probability function expression are as follows:

$$\ln \lambda_i = \ln E_i + \ln X_i \beta_i + \varepsilon_i$$

$$P\left(y_i | \varepsilon_i\right) = \frac{\left[\lambda_i \exp\left(\varepsilon_i\right)\right]^{y_i} \exp\left[-\lambda_i \exp\left(\varepsilon_i\right)\right]}{y_i!}$$

The variance of dependent variable y in Poisson regression is λ_i. In the negative binomial regression model, the variance is equal to $\lambda_i(1+k\ \lambda_i)$, when the value of K approaches 0, the negative binomial regression is almost equal to Poisson regression. $K> 0$ indicates that there is over discretization, which is the key factor in the negative binomial regression model.

5 Data analysis and results

5.1 The first stage: FSI

In the first stage, article title (content perception), push position (environment), and feedback are identified as the essential factors that significantly influence the reading rate, while the new user base has no significant influence on reading rate. The specific analysis is as follows.

The result of the "LR test of alpha = 0" is significant and indicates that we should choose the negative binomial regression model since the overall data in the first stage is over-discrete. The p value provided in Table 4 shows that the read rate of the articles will be positively affected under one of the following two conditions. First, the article title contains emotional symbols. The push position (environment) is the headline, and the overall comment rate and the comment rate of the previous day (Feedback) passed the significance test of 99% confidence because the p value is less than 0.01. Second, push time (environment) is at bedtime (21:30–23:30), the article title does not exceed 15 words, and the overall favor rate of WOAs the day before (Feedback) passed the significance test of 95% confidence, because the p value is less than 0.05. Additionally, the article type has a positive impact on the reading rate of the articles.

As the article type is a key factor affecting the reading rate, we further analyze three types of articles—life, entertainment, and information.

We also chose the negative binomial regression model since the negative binomial regression results of life articles indicated the data are too discrete. Table 5 provides the negative binomial regression results of the life articles in the first stage. In 1169 background operating data of life articles, the article title containing emotional symbols, the push position of headline, and the feedback still passed the significance test of 99% confidence, because the p value is less than 0.01. Meanwhile, push time for bedtime also has a significant positive impact on the reading rate. Unlike the regression results of the overall data, the article title not exceeding 15 words no longer had a significant positive impact on the reading rate because the p value is greater than 0.05

Table 4 Negative binomial regression results in the first stage

Construct	Variable	IRR	St. err.	*t* value	*p* value	Sig.
New user base	Proportion of new users on the day	0.724	1.563	−0.15	0.881	
	Proportion of new users on the previous day	0.297	0.543	−0.66	0.507	
Push position	Whether headline	3.048	0.077	43.94	0	***
Push time	Whether weekend	1.025	0.028	0.91	0.364	
	Whether noon break	1.051	0.124	0.42	0.671	
	Whether quitting	1.056	0.035	1.66	0.098	*
	Whether bedtime	0.949	0.036	−6.39	0.035	**
Article type	Life	1	.	.	.	
	Entertainment	1.303	0.056	6.11	0	***
	Information	0.567	0.02	−16.33	0	***
Article title	Emotional	0.874	0.021	−5.55	0	***
	Not more than15	0.941	0.025	−2.27	0.023	**
Feedback	Comment rate	1.52E-09	3.89E-09	−7.94	0	***
	Share rate	0.840	0.492	−0.30	0.766	
	Like rate	4.14E+10	6.14E+10	16.5	0	***
	Favor rate	0.013	0.024	−2.32	0.02	**
	Constant	0.009	0	−97.31	0	***

***, **, and * indicate $p<0.01$, $p<0.05$, and $p<0.1$, respectively

Similarly, in Table 6, among the background operation data of 648 entertainment articles, the article title and title push within 15 words passed the significance test of 99% confidence, because the *p* value is less than 0.01. At the same time, at a 95% confidence level, push time for bedtime positively affects the reading rate of articles because the *p* value is less than 0.05. The number of characters less than 15 words no longer has a significant positive impact on the reading rate of articles.

In Table 7 in the background operation data of 1152 information articles, the new user group, the push position of the title, the bedtime, and the number of article titles within 15 words passed the significance test of 99% confidence. Comparatively, among the factors of feedback effect of the previous day's WOAs, only the overall sharing rate has no positive impact on the reading rate, while the overall comment rate, praise rate, and collection rate pass the significance test.

Table 5 Results of life articles in the first stage

Construct	Variable	IRR	St. err.	*t* value	*p* value	Sig.
New user base	Proportion of new users on the day	4.95E+29	1.94E+31	1.75	0.081	*
	Proportion of new users on the previous day	6.73E+17	2.31E+19	1.2	0.231	
Push position	Whether headline	2.401	0.09	23.33	0	***
Push time	Whether weekend	1.042	0.041	1.03	0.303	
	Whether noon break	1.046	0.176	0.27	0.788	
	Whether quitting	0.934	0.046	−1.37	0.171	
	Whether bedtime	0.651	0.035	−7.97	0	***
Article title	Emotional	0.901	0.032	−2.92	0.003	***
	Not more than 15	0.993	0.043	−0.16	0.874	
Feedback	Comment rate	2.71E-10	8.83E-10	−6.76	0	***
	Share rate	0.613	0.439	−0.69	0.494	
	Like rate	2.27E+10	3.75E+10	14.47	0	***
	Favor rate	0.288	0.759	−0.47	0.637	
	Constant	0.009	0.001	−80.5	0	***

*** and * indicate $p<0.01$ and $p<0.1$, respectively

Table 6 Results of entertainment articles in the first stage

Construct	Variable	IRR	St. err.	*t* value	*p* value	Sig.
New user base	Proportion of new users on the day	0.439	0.729	−0.50	0.620	
	Proportion of new users on the previous day	0.089	0.138	−1.65	0.098	*
Push position	Whether headline	2.376	0.097	21.28	0	***
Push time	Whether weekend	0.924	0.043	−1.7	0.090	*
	Whether noon break	1.262	0.207	1.42	0.156	
	Whether quitting	1.092	0.068	1.42	0.156	
	Whether bedtime	0.854	0.067	−2.02	0.044	**
Article title	Emotional	0.966	0.042	−0.79	0.431	
	Not more than15	0.862	0.039	−3.32	0.001	***
Feedback	Comment rate	1.49E+13	1.31E+14	3.43	0.001	***
	Share rate	0	0	−4.37	0	***
	Like rate	2.276	17.038	0.11	0.912	
	Favor rate	0	0	−2.99	0.003	***
	Constant	0.016	0.001	−51.38	0	***

Table 7 Results of information articles in the first stage

Construct	Variable	IRR	St. err.	*t* value	*p* value	Sig.
New user base	Proportion of new users on the day	3.8E+279	6.9E+281	3.56	0	***
	Proportion of new users on the previous day	8.5E-54	1.43E-51	−0.73	0.467	
Push position	Whether headline	4.643	0.204	34.87	0	***
Push time	Whether weekend	1.068	0.049	1.43	0.152	
	Whether noon break	0.710	0.171	−1.42	0.155	
	Whether quitting	0.946	0.064	−0.83	0.407	

Table 8 Negative binomial regression results in the second stage

Construct	Variable	IRR	St. err.	*t* value	*p* value	Sig.
Article type	Life	1	.	.	.	
	Entertainment	0.394	0.018	−20.59	0	***
	Information	1.254	0.05	5.7	0	***
Content relevance	High relevance	1.122	0.038	3.43	0.001	***
	Middle relevance	0.91	0.052	−1.66	0.097	*
Content vividness	Whether include video	1.564	0.073	9.64	0	***
	Picture text ratio	1	0	1.31	0.191	
Feedback	Comment rate	2079.668	14306.136	1.11	0.267	
	Like rate	0	0	−7.5	0	***
	Favor rate	1.16E+20	2.44E+20	21.95	0	***

*** and * indicate $p<0.01$ and $p<0.1$, respectively

5.2 The second stage: PS

In the second stage, content vividness, content relevance, and feedback are identified as the essential factors. Specifically, the content, including multimedia and the high correlation between content and its title, significantly affects the sharing rate. For life articles, content relevance is a nonsignificant factor. The specific analysis is as follows.

"LR test of alpha = 0" is still significant; therefore, the overall data in the second stage is excessively discrete. We continue to use the negative binomial regression model. According to the IRR value shown in Table 8, the article type passed the significance test with 99% confidence, indicating the significant impact on the share rate of articles. The information type of articles has the greatest impact on

Table 9 Results of life articles in the second stage

Construct	Variable	IRR	St. err.	*t* value	*p* value	Sig.
Content relevance	High relevance	1.081	0.064	1.33	0.183	
	Middle relevance	0.863	0.081	−1.57	0.116	
Content vividness	Whether include video	1.357	0.094	4.39	0	***
	Picture text ratio	1.001	0	5.4	0	***
Feedback	Comment rate	3.80E+16	4.98E+17	2.91	0.004	***
	Like rate	0	0	−6.72	0	***
	Favor rate	2.20E+25	8.10E+25	15.84	0	***

***indicates $p<0.01$

sharing, followed by the entertainment type, because the *p* value is the same, and the *t* value of information articles is greater.

From the dimension of the content vividness, it can be seen from the *p* value, at 99% confidence level that video and multimedia information have a significant positive impact on the sharing rate, while the ratio of text to picture failed to pass the significance test. Regarding the dimension of content relevance, the high correlation between content and its title positively affects the sharing rate at a 99% confidence level, while the moderate correlation between content and its title positively affects the sharing rate at a 90% confidence level.

In the feedback effect dimension of information dissemination, both the like rate and the favor rate passed the significance test of 99% confidence, which had a significant positive impact on the sharing rate. By contrast, the related impact of comment rate was not significant. However, by comparing the IRR values of the two independent variables, the positive unit effect of the favor rate is much higher than the positive unit effect of the like rate.

Similarly, we analyzed three types of articles in the second stage. Table 9 provides the negative binomial regression results of the life articles in the second stage. From the dimension of the content vividness, at 99% confidence, the content of articles, including video multimedia information and the text-to-picture ratio, has a significant positive impact on the article sharing rate. From the dimension of the content relevance, unlike the regression results of the overall data in the second stage, the two independent variables (high correlation, moderate correlation) of life articles failed to pass the significance test. Regarding the feedback effect of life articles, besides the like rate and favor rate of push articles passing the prominence test of 99% confidence, the comment rate also passed the prominence test. By comparing the IRR values of three independent variables, we found that the positive unit effect of the favor rate was higher than that of the comment rate, which in turn was much higher than that of the like rate.

It can be seen from Table 10 that for entertainment articles, the positive effect of the article containing video multimedia information (IRR=2.018) is higher than that of the text-to-picture ratio (IRR=0.999) in the content vividness dimension. In the content relevance dimension, high correlation and moderate correlation have a significant positive impact on the article share rate. High correlation (IRR = 1.264) articles have more positive unit influence than moderate correlation (IRR = 0.751). The positive unit influence of the like rate (IRR = 1.830E + 25) of entertainment articles is significantly higher than that of the life articles (IRR = 0.000). Among the three independent variables of the favor rate, comment rate, and like rate of entertainment articles, the positive unit effect of the favor rate is higher than the like rate, and the like rate is higher than the comment rate.

It can be seen from Table 11 that for information articles, the ratio of video multimedia information and text pictures has a significant positive impact on the article's sharing rate at the confidence level of 99% and 95%, respectively. Only one independent variable with high correlation passed the 90% significance test. At a certain level of confidence, the comment rate, like rate, and favor rate of push articles have

Table 10 Results of entertainment articles in the second stage

Construct	Variable	IRR	St. err.	*t* value	*p* value	Sig.
Content relevance	High relevance	1.264	0.095	3.12	0.002	***
	Middle relevance	0.751	0.089	−2.43	0.015	**
Content vividness	Whether include video	2.018	0.180	7.85	0	***
	Picture text ratio	0.999	0	−3.34	0.001	***
Feedback	Comment rate	9.55E+11	1.71E+13	1.54	0.124	
	Like rate	1.830E+25	1.48E+26	7.2	0	***
	Favor rate	6.61E+55	7.47E+56	11.38	0	***

Table 11 Results of information articles in the second stage

Construct	Variable	IRR	St. err.	*t* value	*p* value	Sig.
Content relevance	High relevance	1.073	0.040	1.87	0.062	*
	Middle relevance	1.022	0.072	0.31	0.753	
Content vividness	Whether include video	1.238	0.087	3.04	0.002	***
	Picture text ratio	1.000	0	−2.24	0.025	**
Feedback	Comment rate	3.15E-07	2.03E-06	−2.32	0.020	**
	Like rate	1.46E+15	7.12E+15	7.15	0	***
	Favor rate	1.55E+11	3.00E+11	13.33	0	***

a positive impact on the share rate. The positive unit effect of the like rate is higher than the favor rate, which is much higher than the comment rate.

6 Discussion and conclusion

6.1 Research findings

This study examined the factors influencing information dissemination on WOAs. We identified two stages of information dissemination: "feedback-sympathize-identify" (FSI) and "participate-share and spread" (PS). This study also identified several essential factors and their relative importance to examine their effect in each stage. Several interesting findings emerge from this research.

First, in the first stage, both environment and content perception significantly impact the article's reading rate, while new user base has no significant effects. Specifically, the push position of the headlines and emotional article title are important factors to attract users to read. However, the new user base of information and life articles have a significant positive impact on the reading rate under the condition of obtaining effective information [41, 42]. In the content perception dimension, the article title is a typical external feature, and the length of the title affects the click rate to a certain extent. At the same time, the article title contains emotional punctuation ("?" and "!"); it significantly affects the reading rate of life and information articles. That is, the more emotional the title is, the easier it attracts users to read, which is consistent with the findings of [21]. It shows that emotional expression is associated with higher consumer engagement with a message [43].

Article types also play a vital role in information dissemination. Different article types have different effects on users' information dissemination behavior. In the first stage, the positive effects of the three types of articles are entertainment, life, and information from high to low. Conversely, in the second stage, the order is information, life, and entertainment. The difference between the two stages is reasonable because the depth of user participation in the PS phase is higher than that in the FSI phase. In the first stage, users tend to pay attention to the participation behavior of reading. Therefore, entertainment information can better meet their reading needs to achieve relaxation. In the second stage, users tend to consider the information value reflected in the article when sharing diffusion information with other users. Therefore, the promotion effect of information type on article sharing is the main in the second stage.

In the second stage, content judgment has a significant impact on the article sharing rate. Multimedia information significantly improves the sharing rate of the three types of articles (information, life, and entertainment), indicating the importance of the vividness for the article. Researchers have pointed out that the matching of user's favor behavior and needs provide a resource basis for information sharing in the future [44]. Combined with the significance level and IRR value, the sharing rate of three articles is positively influenced by the favor rate of articles. Additionally, content relevance between title and content also affects the sharing rate of the three types to a large extent. It implies that the entertainment articles cannot only rely on exaggerated headlines to attract people's attention but also pay attention to the influence of content relevance. Otherwise, it may result from negative effects on the depth of dissemination of articles.

Finally, the two stages interact with each other. The behavior of the day (like, comment, collection, and sharing) reflects the popularity, practicability, and value of the article, which will have a significant impact on the reading rate of the next day and the sharing rate of the day. Specifically, for life articles, the rate of likes has the greatest impact on the reading rate, and likes, comments, and collections have a great impact on the sharing rate. For entertainment articles, the review rate, sharing rate, and collection rate have a great impact on the reading rate, and the likes and collections have a great impact on the sharing rate. For information articles, the praise rate has a certain impact on the reading rate, praise and collection have a great impact on the sharing rate, and comments have a certain impact on the sharing rate.

6.2 Theoretical implications

This study contributes to the existing literature in several ways. First, considering the characteristics of user behavior on mobile social media, it proposes a new model called FSIPS by adding the feedback effect into the process of information dissemination. Many studies in the information dissemination field focus on sharing behavior (e.g., retweeting, forwarding) [4, 8, 15, 19]. With the rapid growth of social media, users can apply more interactive functions, such as like, comment, and favor, leading to various behaviors affecting factors that are neglected in existing research. This study has explored these factors and added them into the model as factors of feedback. The experimental results show that the feedback effect has a significantly positive impact on the article reading rate and sharing rate.

Second, existing studies analyzed information dissemination behaviors from only two limited levels of users and content [15, 34]. This research has investigated information dissemination behaviors from novel perspectives (i.e., environment, feedback) beyond users and content. It has proposed a comprehensive model to grasp a holistic picture of the information dissemination process. Considering the environmental factors, the constructs validated from this study are more appropriate to the characteristics and the context of social media. As the information diffused at an exponential rate based on fast-changing technology, this study may provide empirical evidence for scholars to better understand the behavior of users and the law of information dissemination on social media.

6.3 Practical implications

Our study also generates new insights for practical operations and provides several suggestions for operators. First, different types of accounts should develop appropriate operation strategies. Taking the WOAs of entertainment articles as an example, the results show that WOAs operators cannot only rely on interesting, entertaining, and exaggerated headlines to attract people's attention but also on the relevance between titles and contents. The excessive deviation between the title and content affects the in-depth participation and sharing diffusion behavior of users. Second, followers play an important role in organizations' social media strategies [20], and the interaction between followers and accounts is conducive to the development of accounts. Focusing on users' needs at different stages in information dissemination is critical. In the FSI stage, the interest in information can stimulate user attention to a certain extent and satisfy users' self-relaxing needs. The information value of WOAs articles is an important factor in promoting users' in-depth participation behavior. Therefore, in the PS stage, operators should stress the content value and realize the information sharing and diffusion. This study will also provide implications for social media companies to improve their products during the COVID-19-influenced periods where the social media are widely used.

6.4 Limitations and future research

In this study, the authenticity and validity of experimental data and experimental methods are guaranteed. However, this study still has some limitations that should be addressed. First, our data were collected from a social media platform in China, and so the results may be influenced by the national culture and platform context to some extent. Owing to the low posting frequency of some WOAs, there will be errors in the feedback effect variables in the model.

Future studies should conduct a cross-cultural investigation on this subject to obtain a better understanding of the effect of information dissemination from a global perspective. Furthermore, other suitable indicators can be considered in the FSIPS model. Future research can also explore a new algorithm for the matching degree between title and content keywords and gradually improve the efficiency of judging the relevance between title and content.

Acknowledgements We would like to thank National Natural Science Foundation of China (NO. 72061147005) and School of Interdisciplinary Studies in Renmin University of China and Metaverse Research Center in Renmin University of China for providing support for this research.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Cao X, Gong M, Yu L, Dai B (2020) Exploring the mechanism of social media addiction: an empirical study from WeChat users. Internet Res 30(4):1305–1328
2. Tseng FC, Cheng TCE, Li K, Teng CI (2017) How does media richness contribute to customer loyalty to mobile instant messaging? Internet Res 27(3):520–537
3. Han Y, Lappas T, Sabnis G (2020) The importance of interactions between content characteristics and creator characteristics for studying virality in social media. Inf. Syst. Res 31(2):576–588
4. Shi J, Hu P, Lai KK, Chen G (2018) Determinants of users' information dissemination behavior on social networking sites: an elaboration likelihood model perspective. Internet Res 28(2):393–418
5. Zhang W, Du W, Bian Y, Peng CH, Jiang Q (2020) Seeing is not always believing: an exploratory study of clickbait in Wechat. Internet Res 30(3):1043–1058
6. Baird CH, Parasnis G (2011) From social media to social CRM: reinventing the customer relationship. Strat Lead 39(6):27–34
7. Zhang Y, Moe WW, Schweidel DA (2017) Modeling the role of message content and influencers in social media rebroadcasting. Int. J. Mark. Res 34(1):100–119

8. Gu C, Xu YC, Yao M (2020) Effect of appeal orientation in sponsored ad forwarding on a microblog platform. Internet Res 30(6):1835–1856
9. Stieglitz S, Dang-Xuan L (2013) Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. J Manag Inf Syst 29(4):217–248
10. Pöyry E, Parvinen P, Malmivaara T (2013) Can we get from liking to buying? Behavioral differences in hedonic and utilitarian Facebook usage. Electron. Commer. Res. Appl 12(4):224–235
11. Rishika R, Kumar A, Janakiraman R, Bezawada R (2013) The effect of customers' social media participation on customer visit frequency and profitability: an empirical investigation. Inf. Syst. Res 24(1):108–127
12. Liang Y, Kee KF (2018) Developing and validating the A-B-C framework of information diffusion on social media. New Media Soc 20(1):272–292
13. Weigel FK, Hazen BT, Cegielski CG, Hall DJ (2014) Diffusion of innovations and the theory of planned behavior in information systems research: a metaanalysis. Commun. Assoc. Inf sys 34(1):619–636
14. Berger J, Schwartz EM (2011) What drives immediate and ongoing word of mouth. J Marketing Res 48(5):869–880
15. Kim J, Bae J, Hastak M (2018) Emergency information diffusion on online social media during storm Cindy in U.S. Int J Inf Manage 40:153–165
16. Liu Z, Liu L, Li H (2012) Determinants of information retweeting in microblogging. Internet Res 22(4):443–466
17. Zhang L, Peng TQ, Zhang YP, Wang XH, Zhu JJH (2014) Content or context: which matters more in information processing on microblogging sites. Comput. Hum. Behav 31(1):242–249
18. Harvey C, Stewart DB, Ewing MT (2011) Forward or delete: what drives peer-to-peer message propagation across social networks? J. Consum. Behav 10(6):365–372
19. Ho JYC, Dempsey M (2010) Viral marketing: motivations to forward online content. J. Bus. Res 63(9):1000–1006
20. Tang Z, Chen L (2020) An empirical study of brand microblog users' unfollowing motivations: the perspective of push-pull-mooring model. Int J Inf Manage 52:1–11
21. Berger J, Milkman KL (2012) What makes online content viral? J Mark Res 49(2):192–205
22. Zaman T, Fox EB, Bradlow ET (2014) A Bayesian approach for predicting the popularity of tweets. Ann Appl Stat 8(3):1583–1611
23. Cheah JH, Ting H, Cham TH, Memon MA (2019) The effect of selfie promotion and celebrity endorsed advertisement on decision-making processes: a model comparison. Internet Res 29(3):552–577
24. Hsu CL, Liu CC, Lee YD (2010) Effect of commitment and trust towards microblogs on consumer behavioral intention: a relationship marketing perspective. Int. J. Electron. Bus 8:292–303
25. Susarla A, Oh JH, Tan Y (2012) Social networks and the diffusion of user-generated content: evidence from YouTube. Inf Syst Res 23(1):23–41
26. Heath C, Bell C, Sternberg E (2001) Emotional selection in memes: the case of urban legends. J Pers Soc Psychol 81(6):1028–1041
27. Rudat A, Buder J (2015) Making retweeting social: the influence of content and context information on sharing news in twitter. Comput. Hum. Behav 46:75–84
28. Wang XY (2021) Research on news comment WeChat official account title -- data analysis based on SPSS. Hunan Normal University, p 1
29. Dellarocas C, Sutanto J, Calin M, Palme E (2016) Attention allocation in information-rich environments: the case of news aggregators. Manage Sci 62(9):2543–2562
30. Ghose A, Yang S (2009) An empirical analysis of search engine advertising: sponsored search in electronic markets. Mana Sci 55(10):1605–1622
31. Sabate F, Berbegal-Mirabent J, Cañabate A, Lebherz PR (2014) Factors influencing popularity of branded content in Facebook fan pages. Eur. Manag. J 32(6):1001–1011
32. Rathore AK, Kar AK, Ilavarasan PV (2017) Social media analytics: literature review and directions for future research. Decis Anal 14(4):229–249
33. Ray S, Kim SS, Morris JG (2014) The central role of engagement in online communities. Inf. Syst. Res 25(3):528–546
34. Son I, Lee D, Kim Y (2013) Understanding the effect of message content and user identity on information diffusion in online social networks. Paper presented at [PACIS: Pacific Asia Conference on Information Systems], [June 2013], [Jeju Island, Korea].
35. Hassan S, Nadzim SZA, Shiratuddin N (2015) Strategic use of social media for small business based on the AIDA model. Procedia Soc Behav Sci 172:262–269
36. Lee SH, Hoffman KD (2015) Learning the ShamWow: creating infomercials to teach the AIDA model. Mark. Educ. Rev 25(1):9–14
37. Tseng CH, Wei LF (2020) The efficiency of mobile media richness across different stages of online consumer behavior. Int J Inf Manage 50:353–364
38. Li H (2012) Advancing advertising theories and scholarship. Advert Theory:576–582
39. Wei PS, Lu HP (2013) An examination of the celebrity endorsements and online customer reviews influence female consumers' shopping behavior. Comput. Hum. Behav 29(1):193–201
40. Isada F, Isada Y (2014) The influence of trust and empathy in word-of-mouth communication on the internet on consumer behaviour: research based on a questionnaire. Rev Econ Stat 2:108–118
41. Baltas G (2003) A combined segmentation and demand model for store brands. Eur. J. Mark 37(10):1499–1513
42. Dunne Á, Lawlor M, Rowley J (2010) Young people's use of online social networking sites-a uses and gratifications perspective. J. Res. Interact. Mark 4(1):46–58
43. Lee D, Hosanagar K, Nair H (2018) Advertising content and consumer engagement on social media: evidence from Facebook. Manage Sci 64(11):5105–5131
44. Shao S, Li C (2013) Based on the social network evaluation model of short-term interaction with followers micro-blogging marketing. In: 2013 8th International Workshop on Semantic and Social Media Adaptation and Personalization. IEEE, Bayonne, France, pp 8–13. https://doi.org/10.1109/SMAP.2013.7