Research Article

Measurement and Trend Analysis of New Media Coverage Topics Based on Comment Big Data Mining

Meng Shao ⓘ

Media & Communication College, Shandong College of Arts, Jinan 250300, China

Correspondence should be addressed to Meng Shao; z00847@sdca.edu.cn

With the rapid development of Internet, new media news has gradually become the most concerned information source of new media people. New media public opinion is a force that cannot be ignored. It needs monitoring and guidance. The research on hot topic discovery and trend analysis can timely find social hot topics and analyze the trend of topics, which is conducive to grasp the trend of public opinion, so as to correctly guide and maintain social stability. At the same time, the emerging industry of new media has sprung up like mushrooms after a rain, with a rapid momentum. With the advent of the era of big data, the development of new media presents characteristics and advantages that are different from traditional media, but we should also note that the era of big data has both advantages and disadvantages for the development of new media. In this paper, LDA and ARIMA models can be used to calculate and analyze the popularity measurement and trend analysis of new media reports under the background of big data mining. The model designed in this paper has a conclusion: the missed detection rate is reduced by 75.4%. Through experiments, it is found that the accuracy of heat topic detection of the model designed in this paper can reach 84.6%. In trend analysis, the first stage of transmission is called the incubation period. Then after a certain critical point, it will come to the outbreak period. The outbreak period lasted for a period of time, entered a period of plateau, and finally came to a subsidence period.

1. Introduction

Today, we have ushered in the era of big data. That is, data resources have become a common understanding. Obviously, in the future business competition, having massive data will become the development advantage of all enterprises [1]. Detailed to the news industry, it is manifested in the rise of data news. In the daily environment of our work and life, big data has become a hot word of the times [2]. In recent years, the emergence of new media such as Weibo, WeChat, and Douyin has made new media have attracted widespread attention. This series of emerging media, based on data information technology and using new media as a carrier to disseminate information, are jointly building the whole new media industry.

The concept of big data refers to massive, high growth rate, and diversified information assets that require new processing modes to have stronger decision-making power, insight and discovery power, and process optimization ability [3]. Driven by the global big data development wave and the huge industrial big data application demand, big data intelligent analysis technology represented by data mining and machine learning has made great progress [4]. At the level of algorithm, more accurate and effective basic theories and algorithm models of machine learning are constantly emerging, especially the deep learning model represented by the new media of deep nerve has made great success in the fields of image processing, natural language processing, and speech recognition [5]. However, with the development of the Internet and the emergence of various new media, we are often troubled by useless redundant information and repeated and cumbersome information in the process of receiving information [6]. For news audiences, tools that can filter news information in a more intelligent and targeted manner are needed; while for news communicators, it is necessary to increase the popularity of news

and expand the scope of news diffusion. Win and gain in the industry are an advantage in the industry [7]. Nowadays, portal websites such as Netease News, Sina News, Today's Headlines, and Tencent News collect updated news in real time and push hot news to users in time, so that users can know the current hot news in time and accelerate the spread of hot news [8]. With the rapid growth of new media news, news information has gradually become huge and bloated [9]. Due to the virtuality of new media, there are also a lot of junk information, even fraud information in new media. News report is one of the most accessible and concerned new media information types. News report has the characteristics of dynamic change, fast growth, and many related topics [10]. Traditional search engines play an important role in people's new media life. Through traditional search engines, people can retrieve and browse the required information from a large number of new media resources, which greatly facilitates the way of information acquisition.

The news media also influences readers' emotional attitudes and value decisions through the choice of information and the expression of attitudes. At the same time, the interactivity and openness of the Internet provide a channel for people to express their views [11]. For the news released by new media, readers can independently screen the topics they are interested in, and "turn a blind eye" to some news they are not interested in. You can also praise or comment on the news to express your views and attitudes [12]. In fact, under the influence of the big data era and the wave of data news, the mainstream news media in my country have already started a series of data news practices [13]. However, these data news are formed by domestic news practitioners learning and imitating foreign practices, and there are still many problems [14]. In addition, most of these studies on advanced cases are micro-interpretation and sharing of specific measures, which cannot provide a complete production path for data news production of other media and organizations [15]. The research goal of this paper is to use big data mining algorithms to find hot topics from a wide range of new media news reports and to study topic trends. On the other hand, through the hot topic detection and trend research method constructed in this paper, the classified data are modeled, so that the hot topics in each time period and the trend analysis of each topic can be obtained.

2. Methodology

2.1. New Media News Analysis from the Perspective of Big Data.
Big data has brought new opportunities for new media news communication. In terms of new media news, news data has exploded. Finding the value of news data through data analysis and exploration is the focus of current journalists' efforts to meet the reading requirements of the masses. Most news is sudden and temporary. Journalists should have strong adaptability. With the arrival of the era of big data, the media industry has more prominent requirements for journalists' response ability. In the era of big data, journalists in the new media environment should make full use of big data means to collect valuable news clues on the Internet,

and use the beauty and interest of network information to dress up and beautify news pictures and backgrounds, so as to timely deliver high-level and valuable news to news audiences. Today, in the era of big data, the most important resource in news reporting is data, and computers and the Internet are no longer an auxiliary method used by journalists in news reporting [16]. It exerts its great advantages and permeates all aspects of news production. It can not only process and analyze the collection of data in news reports but also assist the discovery, production, and release of news clues. It can be said that it has run through the whole news production process, even including the dissemination and operation of news [17]. As a result, a new news reporting model has emerged, that is, data news. As the research of distributed data parallel computing framework has become the general factual standard of big data processing [18]. Set the relationship table and divide r into R1 and R2, where name, title, gender, and salary are the attribute coefficients in the relationship table R. The data results derived in this paper at R1 and R2 are shown in Tables 1 and 2.

In the era of big data, the new media shows the characteristics different from the traditional media, such as timeliness, information content, and personalization. Especially in the timeliness of information, new media often have more advantages than traditional media [19]. New media relying on new media can often integrate and publish information faster and more conveniently, thus ensuring the timeliness of news or events to a certain extent [20]. In terms of content, new media can accommodate and disseminate a relatively large amount of information. Compared with the limited layout of newspapers and the limited duration of television, the amount of information that new media can disseminate is far from what the communication media can match [21]. On the unique personalized recommendation, the new media has enough right to speak. In the era of big data, new media puts the user's needs in a more important position, analyzes the user's browsing records and other data, tailors-related services for users, and pushes the matching content, for example, the TikTok personalized push [22]. The reporting method of accurate news focuses on the accurate analysis and processing of events in society through the method of social statistics, stripping out the law and significance behind the general phenomenon, and showing it to the public [23]. The core understanding of it lies in obtaining information and reporting facts through the means of social science research [24]. Through the study of multivariate data sets, to explore the main factors affecting news flow, we can accurately predict news popularity, so as to have a better understanding of how to optimize news and more effectively enhance news competitiveness [25]. The heat forecast will also involve the strengthening of decision-making, which will make the news delivery effect more ideal. Once again, a data tree is introduced to represent this form, and the node dependency function in relation table R is shown in Figure 1.

2.2. Research and Text Representation of Hot News.
The text preprocessing stage is the process of transforming text data into structured data. Firstly, this paper uses Python new

TABLE 1: R_1 in node 1.

Name	Title	Gender	Salary
A	DMTS	F	60k
B	DMTS	M	70k
C	MTS	M	60k

TABLE 2: R_2 in node 2.

Name	Title	Gender	Salary
A_1	VP	M	100k
B_1	VP	F	50k
C_1	MTS	F	40k

media crawler technology to form a data set of the crawled new media news, then Chinese word segmentation is carried out on the text, the text is cut into scattered independent word sets, part of speech tagging is carried out on each word, and a stop vocabulary is constructed to remove irrelevant interference words [26]. News text has no fixed structure, no uniform format, and cannot be directly identified. In order to facilitate subsequent processing, the news text needs to be vectorized. After the text is vectorized, each news text can be represented as a vector, and the vector can be used for text similarity comparison, text clustering, classification, etc. [27]. Table 3 shows the general data of obtaining text.

At this point, an LDA-based text representation model can be introduced, which is a common topic model. Let D be the document collection, d be the document, K be the collection of potential topics, z be the potential topic, and w_{di} be the first d of the document ith word, z_{di} is the topic of document d containing the ith word. α, β is the basic distribution of LDA model. ω_d is the polynomial distribution of

topics on a document in the document set. The basic model diagram of general LDA is shown in Figure 2.

The hot news topics discussed in this paper generally refer to the event topics that have become the focus of the society and caused certain influence in the news environment, which are caused by certain events and widely participated by netizens [28]. Generally, the hot topics of news are divided into two types. One is news group events. In this case, it refers to those aggregated groups temporarily composed of some groups or unspecified groups of people caused by certain social contradictions and social consciousness. The second is general news events and entertainment events. This type of new media hot topic is also a new media hot topic highly concerned and participated by Internet users. Such hot topic events are often related to people's general demand for new media information. Such new media hot topics are not obvious and have a negative impact on the new media order. The correctness of distributed function dependency discovery can be solved by data redistribution. The number of data redistributions is equal to the number of attributes.

By observation, most of the functional dependencies can be found after the first few data redistributions. Take the relational table $R = \{A_1, A_2, \ldots, A_m\}$ as an example. For property $A_i, i \in [1, m]$, the set of candidate functional dependencies grouped by A_i can be represented as $G_c(A_i)$

$$G_c(A_i) = \{cFD | A_i \in lhs(cFD), A_j \notin lhs(cFD), \quad \forall j < i\},$$
(1)

where cFD is the dependency of candidate functions. Make $N_c(A_i)$ the size of $G_c(A_i)$. Take the first attribute A_1 as an example:

$$C_{m-1}^1 + C_{m-1}^1 C_{m-1-1}^1 + \cdots + C_{m-1}^{m-2} C_{m-1-(m-2)}^1 = (m-1) \times 2^{m-2}.$$
(2)

After the first data redistribution, all the candidate function dependencies with A_1 in the left will be verified one by one. Therefore, $N_c(A_2)$ is equal to

$$C_{m-1}^1 + C_{m-1}^1 C_{m-1-1}^1 + \ldots + C_{m-1}^{m-2} C_{m-1-(m-2)}^1 = (m-2) \times 2^{m-3} + 2^{m-2}.$$
(3)

Suppose T is the total number of all candidate functional dependencies in the property grid. Generally, $CRF(k)$ can be expressed as

$$CRF(k) = \frac{\sum_{i=1}^{k} N_c(A_i)}{T}.$$
(4)

Figure 3 shows the CRF corresponding to the relationship table containing five attribute groups. As you can see, most of the candidate function dependencies are clustered in the first few attributes. The CRF of the first two attributes is as high as 75%. Therefore, after previous data redistribution, most functional dependencies can be found.
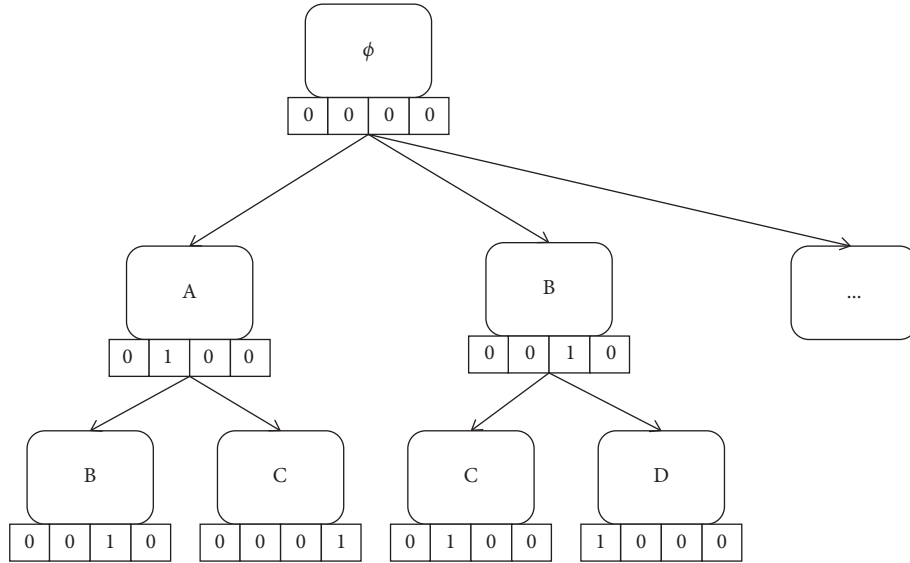
FIGURE 1: Data tree example.

TABLE 3: Acquisition of news data.

Category	Number of news
Domestic	6541
Foreign	5247
Technology	6685
Entertainment	6732
Finance	6428

Time is generally introduced into news reports. In statistical research, a set of random variables arranged in chronological order is used to represent a time series of random events. Time series has a wide range of applications and has many practical applications in the fields of natural science and engineering technology. It is generally defined as an ordered set composed of recording time and recording value:

$$r_t, r_{t+1}, r_{t+2}, r_{t+3}, \ldots \ldots \tag{5}$$

Among them, r represents the value of the record, t represents the time, and the element r_t represents the record value of the time t is r. ARIMA model is a modeling method used to study time series. It is used to describe the linearity of data series and is often used for short-term modeling and forecasting. The model can be subdivided into three categories: AR model, MA model, and ARMA model. The AR model can be expressed as a multiorder autoregressive model, abbreviated as $AR(p)$:

$$x_t = \sum_{i=1}^{p} \varphi_i r_{t-i} + \varepsilon_t, \tag{6}$$

where φ_i is the coefficient of autoregressive term, p is the order of autoregressive term, and ε_i is the random interference sequence. MA model is generally expressed as q-order moving average model, which is noted as $MA(q)$:

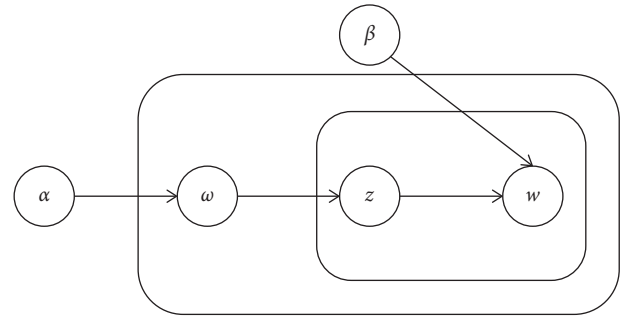$$x_t = \sum_{i=1}^{q} \theta_i \varepsilon_{t-j} + \mu + \varepsilon_t, \tag{7}$$

FIGURE 2: Basic model diagram of topic generation on LDA.

where θ_i is the coefficient of the moving average term, q is the order of the moving average term, and ε_t is the random interference sequence. μ is the initial value. Amra model is an autoregressive moving average model, which is recorded as $AMRA(p, q)$. At this time, the expression is

$$x_t = \phi_0 + \sum_{i=1}^{p} \varphi_i r_{t-i} + \varepsilon_t \sum_{i=1}^{q} \theta_i \varepsilon_{t-j}, \tag{8}$$

where ϕ_0 is the undetermined coefficient, ε_t is the independent error term, φ_i, θ_i is the coefficient of autoregressive term and moving average term respectively, and p, q is the order of autoregressive term and moving average term, respectively.

2.3. Heat Measurement and Trend Analysis. When measuring hot news, the news text needs to be represented first. Since the news text cannot be directly used for calculation by the computer, before topic detection, the news text should be represented in a form that the computer can calculate. The representation of news text is the primary problem to be solved in the preprocessing stage of topic detection, that is, to solve the problem of formatted representation of text in the system for easy recognition and processing. Language
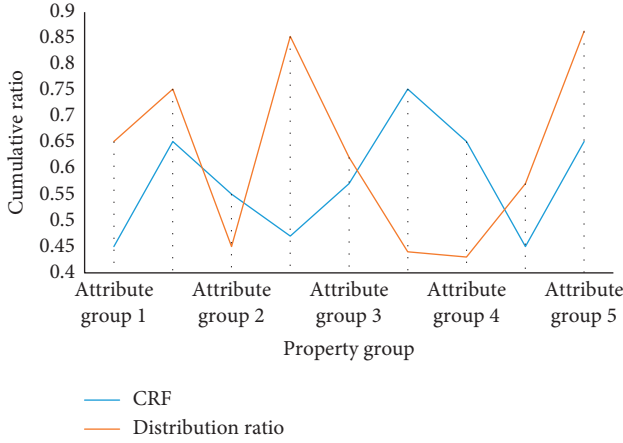
FIGURE 3: Corresponding CRF in the relationship table containing multiple attribute groups.

model is a statistical model based on probability, which can also be regarded as a statistical model for generating a certain language text. In most statistical language models, $n-1$ Markov model is usually used to construct a sentence model. The language model in news reports is constructed as follows. Assuming that the probabilities of words w_i appearing in news reports are independent of each other, the probability that a report s is related to topic tp is

$$P(tp|s) = \frac{P(tp) \times P(s|tp)}{P(s)}$$

$$\approx P(tp) \prod_n \frac{P(w_i|tp)}{P(w_i)}, \tag{9}$$

where n represents the number of words in the report, $P(w_i|tp)$ represents the distribution probability of word w_i in the corpus, and $P(s|tp)$ represents the generation probability of words in topic tp. $w(t,d)$ indicates the weight of the word t_i in the document d, then:

$$idf(t_i, d) = \log\left(\frac{N}{df(t_i)} + 0.01\right),$$

$$w(t_i, d) = tf(t_i, d) \times idf(t_i, d). \tag{10}$$

Among them, $tf(t_i, d)$ represents the probability of word t_i appearing in document d, $idf(t_i, d)$ represents the inverse document frequency of word t_i in the document, N represents the total number of documents, and $df(t_i)$ represents the number of documents containing word t_i. In judging the relevance of news reports and hot topics, the concept of similarity needs to be introduced, which is generally a language-based similarity model:

$$\text{Sim}(d, tp) + \frac{1}{N} \sum_{t \in d} tf(t, d) \log \frac{\gamma P(t|tp) + (1 - \gamma)P(t)}{P(t)}, \tag{11}$$

where d represents the report, tp represents the topic, t represents the feature items, N represents the sum of the feature items in the report d, $P(t|tp)$ represents the probability of t appearing in tp, $P(t)$ represents the probability of

t appearing in the data set, and $tf(t, d)$ represents the word frequency of t in the document. Table 4 is the topic index table of hot topics.

In order to show the start time of each topic and the general development process, topic indices of 4 time nodes are selected in the above table to reflect the complete trend change process of each topic. Each hot topic has experienced life cycle stages such as incubation period, growth period, maturity period, and decline period. When hot topics first appeared, there may be few relevant reports, but with the development of time, the heat increased, gradually became a hot spot, and then reached the peak. From the above analysis, it can be found that the processing process of the first few attributes accounts for a large part of the discovery. Taking this as an example to improve the efficiency of balanced resource utilization, the attributes with low inclination and cardinal number should be given priority. Figure 4 shows the overall flow of Smart FD algorithm.

The basic model functions are linear, exponential, logarithmic, normal distribution, and polynomial functions. Through the analysis of the quantitative indicators of hot news, this paper uses the polynomial function as the model function to carry out regression analysis on the number and proportion of news-related reports and the number and proportion of source new media. The expression of the polynomial function is as follows:

$$y = \sum_{i=0}^{n} \left(\beta_i (x)^i\right). \tag{12}$$

In the fitting of polynomials, the higher the degree of polynomials, that is, the greater the n, the more obvious the fitting degree of curves to real data. Therefore, it is necessary to limit the degree of polynomials:

$$\frac{R_{n+1} - R_n}{R_n} \leq \alpha. \tag{13}$$

The subscript n represents the degree of the polynomial and R_n represents the degree of fit of the n degree polynomial. The rule for selecting the degree of polynomial is: the ratio of the difference between the degree of fit under the n degree and the degree of $n+1$ and the ratio of the degree of fit under the n degree is less than the threshold, that is, when the degree increases, the increase in the degree of fit is not obvious, indicating that the regression function tends to be stable. In this paper, $\text{Trend}_{\text{pred}}$ is used to express the trend prediction result of topic tp and k is defined by using the prediction slope $\text{Trend}_{\text{pred}}$ of trend model under multigranularity.

$$\text{Trend}_{\text{pred}} = k$$

$$= \frac{\sum_{i=1}^{n} \left(k_{\Delta t_i} / \Delta t_i R_{\Delta t_i}^2\right)}{\sum_{i=1}^{n} R_{\Delta t_i}^2}, \tag{14}$$

where n represents the total number of granularity selected in multigranularity fusion, i represents the regression model of the i granularity, and $1 \leq i \leq n$, $R_{\Delta t_i}^2$ represents the fitting degree of the regression model of the i granularity; Δt_i

TABLE 4: Topic index table of hot topics.

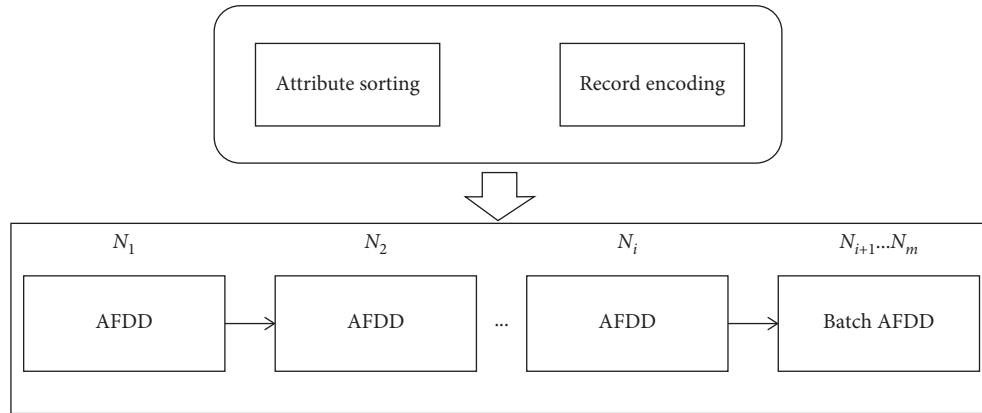|	T1	T2	T3	T4
1	0	253.2	217.4	0
2	124.6	0	158.7	85.3
3	210.3	220.3	0	155.3

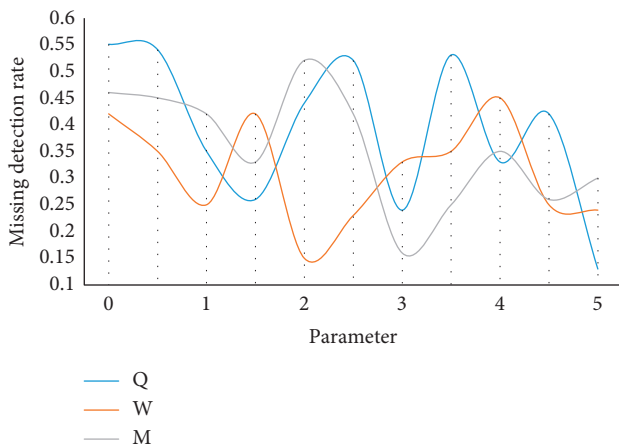FIGURE 4: Overall flowchart of smart FD algorithm.

FIGURE 5: Analysis on the missed detection rate parameter.

represents the value of the i granularity, $\Delta t_i \in (12, 24, 36, 48)$, and $k_{\Delta t_i}$ represents the prediction slope of the prediction point in the regression model of the i granularity.

3. Result Analysis and Discussion

In order to establish a scientific and accurate new media report theme heat measurement and trend analysis model, this paper, on the basis of the above research and analysis, further experimental analysis, hoping to show the effect in specific data experiments. In the hot measurement and trend analysis of whether the detection model can match the topic of new media reports, this paper analyzes the missing rate parameter, the hot topic detection accuracy rate, and the general trend of topic spread. Q represents the missing rate parameter matching the subject reported by the new media. Q represents the hot spot detection accuracy reported by

new media. M represents the overall trend of theme communication. Here are the experimental analysis charts of three different sample sets Q, W, and M in these three indicators, as shown in Figures 5–7.

As can be seen from Figure 5, with the continuous increase of the parameters, the missed detection rate gradually decreases, and finally stabilizes. Based on the experimental results of the four stages of incubation, growth, maturity and decline, the optimal value of the parameter is 2.44, and the missed detection rate at each stage has reached the minimum. Overall, the missed detection rate is decreased by 75.4%. As can be seen from the above figure, the final result of topic detection is that generally, the result of topic detection is worse than that of clustering. In practical application, the report set to be processed is very large, and it is impossible to cluster the data of the whole data set once. Through the experiment, it is found that the accuracy of hot topic detection of the model designed in this paper can reach 84.6%. In the initial stage of the topic, the event news is only spread in a small range, and the object and audience are relatively few. With the gradual spread and expansion of hot topics, there will be some new media or famous people involved at this time. Through their reports and opinions, the whole hot topic explodes at some zero boundary point. The audiences of these new media or well-known figures become the initiators of the next round of communication. Through their forwarding, the cumulative social attention of the whole hot events has increased rapidly. Then the spread of hot topics entered a stable period. At this time, there were new follow-up reports, and the attention of the media and the public continued at a certain level for a period of time. Then, over time, public attention gradually declined and eventually disappeared. Summarizing the above-mentioned transmission process, the first stage of transmission can be called the incubation period. Then, after a certain critical
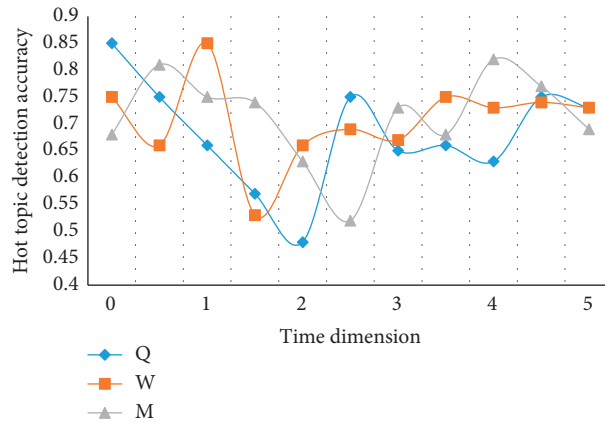
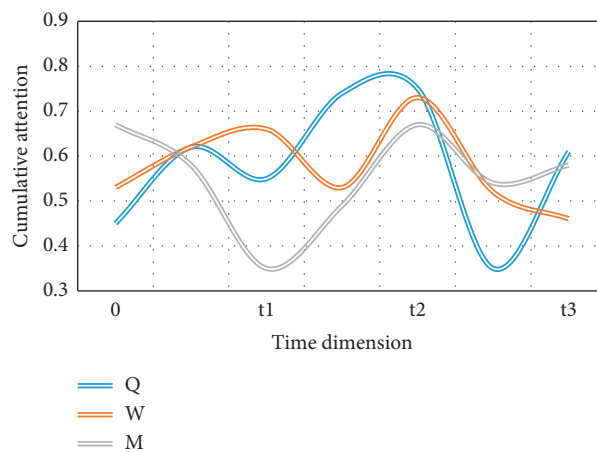FIGURE 6: Analysis of the accuracy of hot topic detection.

FIGURE 7: Analysis of the overall trend of topic dissemination.

point, it will come to the outbreak period. The outbreak period lasted for a period of time, entered a period of plateau, and finally came to a subsidence period.

4. Conclusion

The form of news realizes the transformation of traditional news production, and its advanced nature determines that it is in line with the future development trend of news industry. Compared with traditional news production, it introduces new subjects, puts forward higher requirements for journalists and teams, and promotes the development and transformation of news industry. The progress of the times has not only brought dawn to the development of all industries but also brought new challenges, and journalism is no exception. Therefore, the research on data journalism from the perspective of big data is a very meaningful thing. In this paper, you are proposing an algorithm based on big data to analyze the topic heat measurement and trend of new media reports, and analyze some quantitative characteristics of the topic, including the number of topic-related reports, the number of source websites, and the dispersion of reports. Comprehensively, analyze the indicators of the topic to obtain an energy value of the topic and judge whether a topic is a hot topic according to the energy

value. On the basis of experimental analysis, a conclusion is drawn about the model designed in this paper: overall, the missed detection rate is reduced by 75.4%. Through the experiment, it is found that the accuracy of hot topic detection of the model designed in this paper can reach 84.6%. In trend analysis, the first stage of transmission is called incubation period. Then, wait until after a critical point and then come to the outbreak period. The outbreak period lasts for a period of time, then enters a stable period, and finally comes to the regression period. This paper uses polynomial function as model function, which lacks regression analysis of the number and proportion of news reports. Therefore, future research needs to be combined with more data sources for analysis.

Data Availability

The data used to support the findings of this study are available from the author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] C. Wells and K. Thorson, "Combining big data and survey techniques to model effects of political content flows in facebook," *Social Science Computer Review*, vol. 35, no. 1, pp. 33–52, 2017.

[2] T. A. Maniou, "Semantic analysis of cultural heritage news propagation in social media: assessing the role of media and journalists in the era of big data," *Sustainability*, vol. 13, no. 1, p. 341, 2021.

[3] L. Wang, S. Li, and W. Wang, "A bank liquidity multilayer network based on media emotion," *The European Physical Journal B*, vol. 94, no. 2, pp. 1–23, 2021.

[4] G. Birz, "Stale economic news, media and the stock market," *Journal of Economic Psychology*, vol. 61, no. 3, pp. 384–412, 2017.

[5] L. Gan, S. Li, and Z. Shu, "Big data metrics: time sensitivity analysis of multimedia news," *Journal of Intelligent and Fuzzy Systems*, vol. 38, no. 1, pp. 1–8, 2019.

[6] M. Asgari-Chenaghlu, M. R. Feizi-Derakhshi, L. Farzinvash, M. A. Balafar, and C. Motamed, "Topic detection and tracking techniques on twitter: a systematic review," *Complexity*, vol. 2021, no. 4, 15 pages, Article ID 8833084, 2021.

[7] J. Dong, W. Dai, and J. Li, "Exploring the linear and nonlinear causality between internet big data and stock markets," *Journal of Systems Science and Complexity*, vol. 33, no. 3, pp. 783–798, 2020.

[8] S. Lecheler and C. H. de Vreese, "News media, knowledge, and political interest: evidence of a dual role from a field experiment," *Journal of Communication*, vol. 67, no. 4, pp. 545–564, 2017.

[9] T. G. van der Meer and M. Hameleers, "Fighting biased news diets: using news media literacy interventions to stimulate online cross-cutting media exposure patterns," *New Media & Society*, vol. 23, no. 11, pp. 3156–3178, 2021.

[10] D. K. Kilgo, "When is the "racist" designation truly applicable? News media's contribution to the debatability of racism," *Television & New Media*, vol. 22, no. 2, pp. 186–204, 2021.

[11] R. Braun, "Bloodlines: National border crossings and antisemitism in Weimar Germany," *American Sociological Review*, vol. 87, no. 2, pp. 202–236, 2022.

[12] Y. Feng, P. Zhou, J. Xu, S. Ji, and D. Wu, "Video big data retrieval over media cloud: a context-aware online learning approach," *IEEE Transactions on Multimedia*, vol. 21, no. 7, pp. 1762–1777, 2019.

[13] Y. . Zheng, "Key technologies of media data in-depth analysis system based on artificial intelligence-based big data," *Mobile Information Systems*, vol. 2021, no. 3, Article ID 7191567, 10 pages, 2021.

[14] M. BalaAnand, N. Karthikeyan, and S. Karthik, "Envisioning social media information for big data using big vision schemes in wireless environment," *Wireless Personal Communications*, vol. 109, no. 2, pp. 777–796, 2019.

[15] B. H. Al-Qarni, A. Almogren, and M. M. Hassan, "An efficient networking protocol for internet of things to handle multimedia big data," *Multimedia Tools and Applications*, vol. 78, no. 21, pp. 30039–30056, 2018.

[16] I. I, Q. Sun, Z. Liu, S. Zhang, and S. Han, "The big-data-driven intelligent wireless network: architecture, use cases, solutions, and future trends," *IEEE Vehicular Technology Magazine*, vol. 12, no. 4, pp. 20–29, 2017.

[17] N. A. G. Fredheim, "Life, death or drugs: actor framing power on the news media coverage of health care policy," *European Journal of Communication*, vol. 36, no. 3, pp. 270–285, 2021.

[18] I. Bernetti, G. Chirici, and S. Sacchelli, "Big data and evaluation of cultural ecosystem services: an analysis based on geotagged photographs from social media in Tuscan forest (Italy)," *iForest: Biogeosciences and Forestry*, vol. 12, no. 1, pp. 98–105, 2019.

[19] C. Ordonez and L. Bellatreche, "Guest editorial-DaWaK 2018 special issue-trends in big data analytics," *Data & Knowledge Engineering*, vol. 126, no. 7, Article ID 101730, 2020.

[20] D. Khajeheian and P. Ebrahimi, "Media branding and value co-creation: effect of user participation in social media of newsmedia on attitudinal and behavioural loyalty," *European Journal of International Management*, vol. 1, no. 1, 2020.

[21] S. Karimi, A. Shakery, and R. Verma, "Online news media website ranking using user-generated content," *Journal of Information Science*, vol. 47, no. 3, pp. 340–358, 2021.

[22] G. Cooper, L. Blumell, and M. Bunce, "Beyond the "refugee crisis": how the UK news media represent asylum seekers across national boundaries," *International Communication Gazette*, vol. 83, no. 3, pp. 195–216, 2021.

[23] A. Lazzari, J. M. Pou, C. Dubois, and L. Leblond, "Smart metrology: the importance of metrology of decisions in the big data era," *IEEE Instrumentation and Measurement Magazine*, vol. 20, no. 6, pp. 22–29, 2017.

[24] C. Wang, "Analyzing the effects of cross-border E-commerce industry transfer using big data," *Mobile Information Systems*, vol. 2021, no. 8, 12 pages, Article ID 9916304, 2021.

[25] W. Xu, F. Murphy, X. Xu, and W. Xing, "Dynamic communication and perception of cyber risk: evidence from big data in media," *Computers in Human Behavior*, vol. 122, no. 3, Article ID 106851, 2021.

[26] N. Charron and P. Annoni, "What is the influence of news media on people's perception of corruption? Parametric and non-parametric approaches," *Social Indicators Research*, vol. 153, no. 3, pp. 1139–1165, 2021.

[27] K. Kortum, "Guiding and deploying transformational technologies: TRB hosts partners in research symposium," *TR news*, vol. 307, pp. 38-39, 2017.

[28] M. Rizwan and W. Wan, "Big data analysis to observe check-in behavior using location-based social media data," *Information*, vol. 9, no. 10, p. 257, 2018.