Analysis of Hot Topics and Evolution of Network Public Opinion Based on Social Networks

Lifang Fu[1], Zhuo Chen[2*], Ruiyu Zhang[2], Shuai Liu[2]
[1]School of Science, Northeast Agricultural University, Harbin, Heilongjiang, China
[2]School of Engineering, Northeast Agricultural University, Harbin, Heilongjiang, China,
*Corresponding Author

Abstract:

There are a large number of public opinion information related to the safety of agricultural products on all kinds of network platforms, so it is necessary to quickly and accurately identify the focus and hot topics of public opinion. This paper proposed a method to extract and classify the hot topics of public opinion according to the dependency syntax rules in Chinese sentences. Then, the social network method is used to discover and visualize the changing of different hot topics. Based on the characteristics of the number of nodes and node degree of social networks, this paper analyzed the evolution process of some hot topics combined with different developing stages of public opinion. The result of case study shows that the method of this paper could extract reasonable hot topics from a large number of short text messages, and give a visually show of the community division and evolution characteristics of different topics. It provides effective analysis methods and techniques for the analysis of agricultural network public opinion.

Keywords: *Network public opinion, Hot topic discovery, Dependency syntax, Social network; Hot topic evolution.*

I. INTRODUCTION

With the rapid development of the Online and the continuous improvement of rural informatization, agricultural public opinion incidents have also occurred frequently, such as the "African Swine Fever" incident in 2019 and the "Salmon COVID-19" incident in 2020, whose online public opinion has brought together a variety of different views from all walks of life in society. Since the safety of agricultural products is closely related to the daily life of the general public, it is more likely to cause continuous attention, worry and anxiety of the public[1]. Failure to timely and accurately get the focus of public opinion and the development and evolution of public opinion as well as timely and effective guidance and prevention and control may cause the outbreak of social public opinion and the agglomeration effect of negative emotions, which will have a negative impact on the parties involved in the event as well as social stability and development. Therefore, it is urgent to improve the ability of agricultural network public opinion monitoring and find out the evolution law of public opinion.

Traditional detection methods of public opinion hot topics, such as the hierarchical CURE algorithm, the partition-based K-means algorithm and some classical algorithms such as Single-Pass, are mostly

computationally complex and sensitive to outliers. In recent years, the methods of improved LDA theme model[2-6] and word co-occurrence analysis[7-9] have been mostly used in research. Although the hot public opinion recognition method based on LDA theme has better effect than vector space model[10], it ignores the sequential order and semantic connection between words and is greatly affected by word frequency. As a result, the production result is relatively single, which cannot reflect the correlation and hierarchy of different hot topics[11], so it is not conducive to the exploration of the evolution law of public opinion events.

In this paper, a social network-based public opinion hot topic mining and visual analysis method was proposed. In the analysis of short texts of public opinion information, dependency parsing rules are applied to make the extracted hot words have clear semantic connection. At the same time, the approval number and forwarding number of published information are included in the calculation of the correlation weight between words, so as to construct a relationship diagram of hot words with clear relationship, easy quantification and prominent key points, and make visual analysis to explore the evolution law of hot topics in different stages of public opinion.

1.1 Introduction to Public Opinions about the Safety of Rural Products on the Internet

In this paper, the empirical analysis and model test were mainly conducted on the online public opinions of the "Salmon COVID-19" incident in June 2020. On June 11, 2020, Beijing reported a confirmed case of COVID-19, which suddenly broke the record of Beijing having no new local cases for 56 consecutive days. Then, on the evening of June 12, COVID-19 was detected from the chopping board of imported salmon during random inspection by relevant departments of Beijing Xinfadi Wholesale Market, which aroused heated discussion and a certain degree of panic among netizens. The daily increase of Weibo blog posts around this event was close to 4,000. The news on the evening of June 16 reported that COVID-19's receptor was ace2 protein, not specific to the lung, but present in many animal tissues, which triggered the second wave of heated discussion among netizens. On June 18, there was a sudden drop in the discussion of the incident. By late June, there had been very little discussion information among the netizens on the incident, indicating that the public opinion incident had basically dissipated. The comments from Weibo on the "Salmon COVID-19" incident were collected and collated.

1.2 Information Acquisition and Data Processing of Public Opinion Events

In this paper, first of all, a total of 17,579 blog posts at Sina Weibo from June 11, 2020 to July 10, 2020, which were related to the "Salmon COVID-19" incident were obtained from the domestic mainstream social platforms by using python web crawler technology. According to the development process of the event and the corresponding daily Weibo statistics, the public opinion event is divided into four stages, namely, infancy stage, fermentation stage, peak stage and dissipation stage, as shown in Fig. 1.
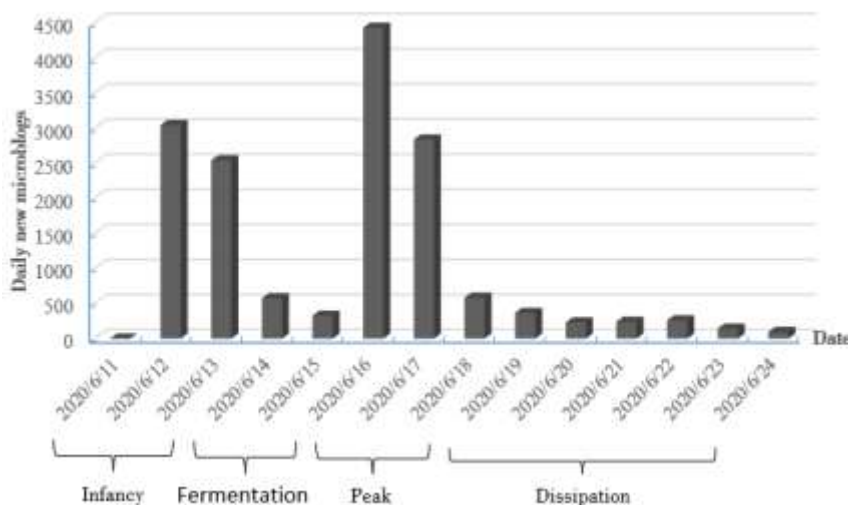
Fig 1: Daily added numbers of Weibo and stage division of 'salmon Covid-19'

Then, the text analysis and topic extraction were carried out for the public opinion information of each period, so as to find out the main focus of netizens in the event, and explore the evolution trend of hot topics in different stages of the public opinion event, in order to timely and accurately understand the public concern, reveal and discover the evolution law of public opinion, and effectively monitor the development of public opinion.

II. MATERIALS AND METHODS

2.1 Dependency Parsing of Hot Words

When discovering and analyzing co-occurrence of words, the method of exhaustion is generally adopted in the traditional multi-word co-occurrence analysis. For example, in a short text of public opinion events, aiming at the sentence "virus was detected from the salmon cutting chopping board", five words are firstly segmented, and 10 associations between words are generated by the method of exhaustion (as shown in Fig. 2(a)), which is obvious that some of the associations between words are redundant and meaningless. Thus it is clear that this method cannot quickly and accurately find the hot topic of this public opinion event. In order to solve the problem that the bag-of-word model ignores the syntactic structure and is difficult to effectively analyze the sentence semantics[12], dependency parsing was used to process the public opinion text information in this paper.

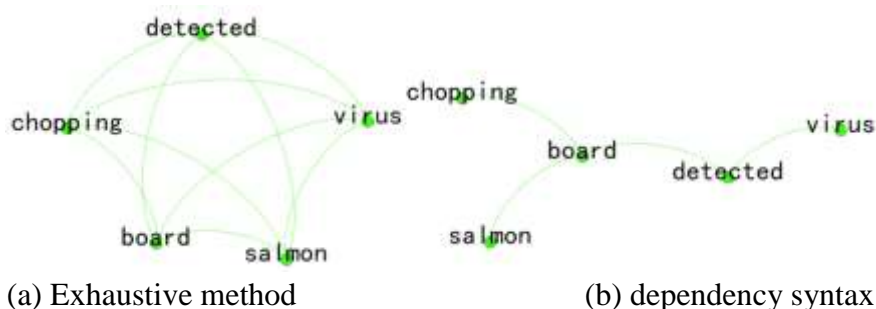(a) Exhaustive method (b) dependency syntax

Fig 2: Comparison chart of word co-occurrence and dependency syntax

Dependency parsing[13], first proposed by French linguist L.Tesniere, analyzes sentences into a dependency syntax tree to describe the dependency relationship between words. It holds that a sentence is an organized unit composed of individual words, and there is a dominant and subordinate relationship between each component of the sentence, among which the dominant word is called Head and the dominated word is called Dependency. According to the previous studies, relational groups that contain the structure in the shape in dependency parsing results are extracted by specifying rules for affective analysis in the reference[14]. According to the corresponding rules in dependency parsing, the subject and the object are related through agent relation and patient relation in reference[15] for social network analysis.

In order to highlight the relationship between words and the overall comprehensibility of sentences, in this paper, the triple structure of < comment words, dependency relations, relation words > was adopted to express the logical differences between texts, which is helpful to guess the true semantics of sentences from the overall perspective. The proposed syntactic rules and examples are shown in Table I, and the dependency relationships are shown in Figs. 3 and 4. The original text of Example 1 is: "Virus was detected on salmon cutting board", and the original text of Example 2 is: "Salmon and related aquatic products were taken off the shelf overnight".

TABLE I. Triad extraction rules and examples

| Rules	Dependence relationship diagram	Triple integration
Rule 1.1: Hot words are subject or object, which are related by agent relation (ER) and patient relation (AR).	E.g.1 as shown in Fig. 3	< salmon, ER-AR, virus >, < chopping board, ER-AR, virus >
Rule 1.2: Hot words are in compound noun modifier relation (CNM). Hot words are attributives or modified words, which relate attributives to modified words.	E.g.1 as shown in Fig. 3	< salmon, CNM, chopping >, < chopping board, CNM, chopping>
Rule 1.3: Hot words relate the subject and predicate when they are in the agent relation (ER) and non-patient relation (AR).	E.g. 2 as shown in Fig. 4	< salmon, ER, off the shelf >
Rule 1.4: If there is a coordinate relation (CR), all words of CR relation act as the same components in the sentence.	E.g. 2 as shown in Fig. 4	< salmon, CR, aquatic products >, < aquatic products, CR, off the shelf >
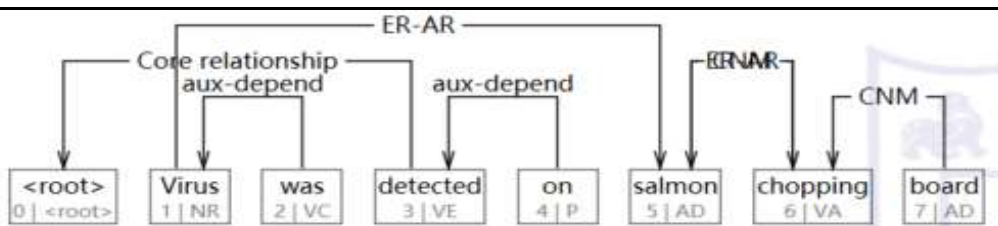
Fig 3: Dependency example 1
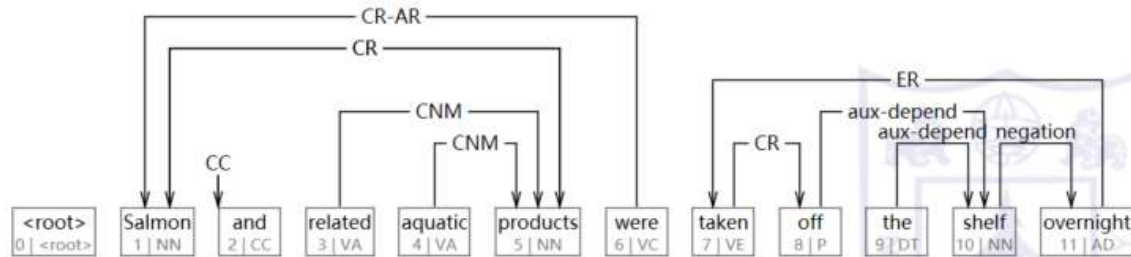
Fig 4: Dependency example 2

The above example sentences were subject to the dependency parsing according to the rules as specified in Table I, and only 4 links were established among 5 participles (as shown in Fig. 2(b)). The figures clearly show that the four links have clear semantic and logical relationships, which indicates that dependency parsing is more conducive to mining the text content of the short text. It also breaks through the constraint of traditional manual dependency relationship formulation, improves the effect of hot word extraction, and thus improves the efficiency of public opinion information analysis.

2.2 Rules for the Integration of Compound Nouns in Information

According to the relationship diagram of hot words obtained by dependency parsing rules, the frequency and weight of each hot word in public opinion information were further counted and analyzed, and its popularity was reflected by dots of different sizes in the relationship diagram[16], and the word co-occurrence relationship diagram of Example 2 was obtained (as shown in Fig. 5). Although the relationship between hot words in this picture is clear, there are many compound noun modifiers, such as "COVID-19" and "virus". In this paper, compound nouns which appeared continuously were formed into compound phrases, which is helpful to understand the overall semantics.

Fig 5: Hot words co-occurrence before interation of example 3

Fig 6: Hot words co-occurrence after interation of example 3

Due to the characteristics of parsing after word splitting in dependency syntax, the complicated interrelation increases the difficulty of understanding. For example, in Example 1, "salmon cutting board" can be used as a nominal phrase, but it appeared only a limited number of times in massive public opinion texts, and any word in the phrase might be associated with a large number of other hot words, so only two-word phrases were extracted to solve this problem. The following integration rules of compound nouns were put forward.

Rules for integration: Phrases that can be integrated according to neighboring words are extracted from N consecutive compound noun modification relationships, and integrated and correlated from high to low according to the times of their appearance in the text data set. Only phrases with the same composition in the sentence before the combination are extracted, and public opinion information texts are searched circularly for integration until the end. The relationship of hot words obtained after processing Example 3 according to this integration rule is shown in Fig. 6. The number of hot words in this example is reduced from 9 to 7, and the connection between words is more concise and effective. In Example 3, the original text is the topic combination of "Covid-19 was detected by the cutting board of salmon" and "the salmon were removed from the shelves by Beijing Xinfadi Merchants".

2.3 Community Division and Weighting of Hot Words

Since the same public opinion event contains different hot topics, in order to detect related objects within the same topic and then obtain hot topics under different topics, the community division and weighting of hot-word relationships were performed in this paper using the modular operation of social relationship network[17].

The concrete measurement standard of the modularity effect of the community network is represented by the modularity[18] Q value, expressed as

$$Q = \sum_i (e_{ii} - a_i^2) \tag{1}$$

Where,

e_{ii}= the ratio of the edge number of module i to the edge number of the whole network;

a_i= the ratio of the node degree of module i to the node degree of the whole network.

According to formula (1), the size of Q is determined by the weight of the edge[19]. The edge weights are generally expressed as the co-occurrence times of a pair of relationship objects. In order to better reflect the focus of netizens' attention, the number of likes and forwards of blog posts were introduced in this paper when calculating the edge weights between nodes. The highly liked and mostly forwarded blog posts generally come from celebrities or self-media with unique opinions, even those directly related to events. In addition, it is easier for netizens to browse and obtain highly liked and mostly forwarded blog posts, which are more influential. The strong influence of opinion leaders or related parties was highlighted by including likes and forwards as part of the calculation of edge weights:

$$w_i = \left\lceil \sum\nolimits_{j=1}^{n} \sqrt{e_{ji}} \sqrt{p_j} \sqrt{s_j} \right\rceil \tag{2}$$

Where,

w_i=the upward rounding of the weight of the i-th relation pair;

n =the total number of responses;

e_{ji}i =the number of occurrences of the i-th relation pair in the j-th response;

p_j=the number of likes of the j-th post;

and s_j= the number of forwards of the j-th post.

After the community division and weighting of the hot words, the social network visual representation of the hot words in Example 3 is shown in Fig. 7, which is a keyword co-occurrence network structure, in which one node represents a keyword, the size of the node reflects its popularity, the connection and thickness between the nodes represent the co-occurrence relationship and intensity between the keywords, and different colors represent different themes. The blog posts in the infancy stage were used as a demonstration, and the data of nodes and edges are shown in Table II and Table III respectively.

Fig 7: Hot words network after community division and weighted

TABLE II. Nodes data of hot words network

Number	Label	Weight	Module
0	Salmon	32.7566	0
1	Beijing	30.9677	1
2	Xin Fadi	14.6628	1
3	Markets	12.5299	2
4	Taken down	11.2249	2
5	Chopping Board	19.2353	0
6	Coronavirus disease	30.5450	0

TABLE III. Edges data of hot words network

source node	target node	weight
0	6	23.0624
3	4	8.3168
0	3	8.6403
1	2	18.9880
0	1	22.3229
5	0	11.4903

2.4 Text Processing of Public Opinion and Extraction of Hot Word Relations

In this paper, 17,579 Weibo blog posts of this public opinion event were preprocessed by using the dependency parsing rules and the relationship division of hot word network mentioned above, so as to reduce the confusion of the relationship network, extract the true and effective triple set, and generate the hot word candidate set, which provides data support for the subsequent visual analysis[20].

First, the data was cleaned out to eliminate invalid data, such as blog posts with only "forwarding" or a large number of special symbols in the text. Then, the jieba toolkit was used to segment the original text set, and the data after segmentation was filtered by the stop word list of Harbin Institute of Technology, and then the word frequency was counted. According to the word frequency, the words with low event

relevance degree were deleted, and some single-word verbs such as "say", "do" and "go" in high-frequency words were artificially processed, so as to screen and retain the words which are truly meaningful to public opinion events. Hot words relation set is the basis of dependency parsing. According to the rules for integration of compound nouns, the corresponding phrases were integrated and imported into the jieba dictionary for secondary word segmentation, and the final hot words candidate data set was obtained after counting the word frequency.

2.5 Network Visualization Analysis of Hot Words in Public Opinion Events

The data of nodes and edges obtained in the aforementioned parsing process were visually displayed in order to more intuitively reflect the hot topic of the public opinion event as a whole. In this paper, Gephi was used as the visualization tool, which is an open source and cross-platform JVM-based complex network analysis software, mainly used in various networks and complex systems. The node degree of the hot words and the weight of the connecting edges extracted from the public opinion text processing were calculated respectively, and the calculation results include the information such as the label of the hot words (nodes), the popularity, the community division, the source of the edges pointing to the nodes, the weight, etc. After the above data information was imported into Gephi for debugging, the most suitable UI layout- OpenOrd was selected.

The key hot words and their relationship network in different stages of the public opinion event have different characteristics and changing rules. For example, in the infancy stage of the event, the overall visual image of public opinion information is shown in Fig. 8, and each topic classification is reflected in different colors.

Fig 8: Overall visualization image

By deconstructing and analyzing the visual images of hot topics, more specific hot topics can be identified. In the figure, nodes of different sizes correspond to hot words of different popularity, and the thickness of the connection between nodes indicates the difference of correlation between hot words, in which there are many related objects in the core nodes, such as salmon, virus, infection and COVID-19, and few related objects in the non-core nodes, which often cluster around the core nodes to form an expansion area. According to the core association shown in the figure, more specific hot topics can be condensed. For example, the strong association between the hot words "salmon", "infection" and "eating" reflects the hot discussion among netizens on topics such as "whether salmon can be infected with the virus" and "whether salmon can be eaten after all". In addition, more specific information related to hot topics can be dug out along the core association and core extended association to the tail node, such as "salmon", "cutting board", "seafood market" and so on, which together reflect more clear problems derived from events.

2.6 Analysis of Hot Topics and Evolution in Different Stages of Public Opinion Events

2.6.1 Hot topics in infancy period

The core nodes of the first hot topic (as shown in Figure 9) include aquatic products, seafood, vendors, supermarkets, and daily food stores. Among them, aquatic products, supermarkets, and seafood are closely related, reflecting the impact of the salmon incident, related aquatic products and seafood sales have been affected, and many supermarkets have removed salmon overnight, and the Beijing-Shenzhen Seafood Market in the expansion area, Loss and investigation are further explanations of the theme, and from the temporary, raw food, fear, masks, and inability to eat, it can be seen that netizens are worried about the salmon incident.
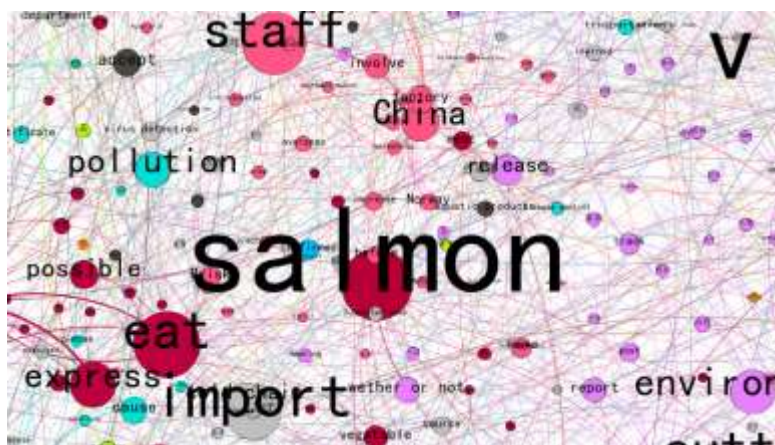

Fig 9: Overall Visualization of infancy stage

The core nodes of the second hot topic are Norway, Beijing, customs, and import. Among them, Norway has the greatest degree of association with imports, which highlights the doubts of netizens about the sanitation of salmon imported from Norway after the incident. The correlation between customs, virus, and testing typically reflects the hot topics of discussion among netizens. The pollution and investigation in the customs expansion area easily reflects the dissatisfaction of netizens with the incomplete inspection of imported products by the customs.

2.6.2 Hot topics in fermentation period

The popularity of public opinion events after the infancy stage has decreased, whose first hot topic is similar to that in the infancy stage, with salmon, COVID-19, chopping board and host as the core nodes. The difference is that the correlation intensity between salmon and host has risen to the first place, while that between salmon and virus has decreased a lot, the concerns of netizens have gradually shifted to deeper issues which shows that with the passage of time. In addition, a number of hot words with high correlation strength with salmon have emerged, such as parasites, mammals, and pollution, which reflect the further information mining by Internet users.
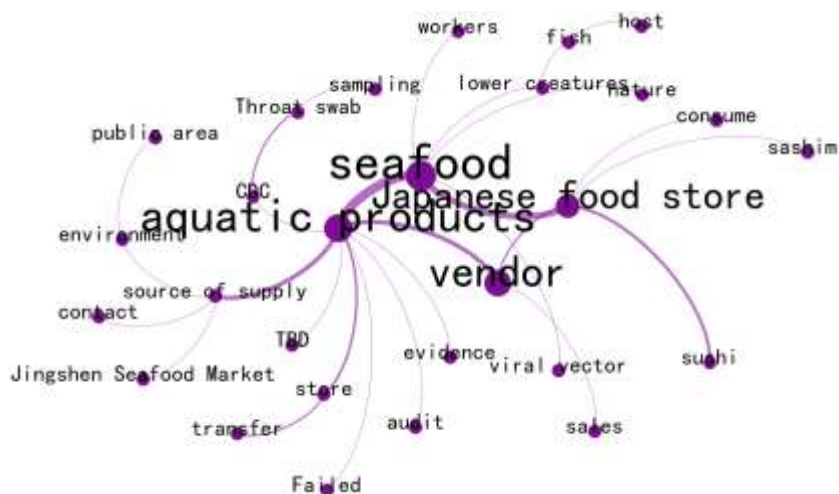
Fig 10: The first hot topic of infancy stage

The core nodes in the second hot topic shrink from the four nodes in the infancy stage to the single node that is spread around the Japanese restaurant. At this time, the discussion focuses of netizens spread downward into a group, and different communities are formed in the image expansion area, which is consistent with the public opinion events.

The features of the third hot topic are similar to those of the third hot topic in the infancy stage, whose core node indicates that there is no significant change in the discussion center of netizens. However, the subtle change in netizens' attitudes can be detected from the change of hot words in the expansion zone. The hot words such as Norway and pollution in the infancy stage seem to convey some netizens' worries about the hygiene of imported salmon from the Europe. The cold chain transportation, regulatory agencies and system corruption in this period reflect that netizens are beginning to realize the problems existing in the country.

2.6.3 Hot topics in the peak stage

As "COVID-19 receptor was exposed to ace2 protein", the salmon incident with a slight decline in popularity was once again pushed to the forefront. The core nodes of the first hot topic were salmon, COVID-19, mammals, ACE2, Zhang Yuxi, garbling and confusing the public, among which virus had become the core word with the highest correlation with salmon. The key information of public opinion events was reflected in the extension area. Netizens discussed that "COVID-19's receptor is ace2 protein" and "Ace2 receptor is not specific to lung and exists in many tissues", which reflected netizens' deep exploration on whether salmon can infect Covid-19 and overturned the previous statement that "how can salmon infect COVID-19 without lung". However, experts pointed out that "COVID-19's known receptors are all mammals, and the probability of salmon carrying virus is almost zero". The finally cleared salmon has gradually reduced the popularity of the topic.

Due to the public punishment for the Beijing-related epidemic by the *Beijing News*, the second topic mainly revolves around the first report of the salmon incident and an interview with Zhang Yuxi, chairman of Xinfadi. He said in an interview that the relevant departments had detected COVID-19 in the chopping board of the cut imported salmon during the sampling inspection, and the source of this product was from Jingshen Seafood Market. In fact, the virus was detected in 40 environmental samples of Xinfadi in the official notification, and Zhang Yuxi only mentioned the salmon chopping board, which caused some confusion. Such terms as prosecution, rumor source and punishment reflect netizens' anger over the losses suffered by seafood vendors and Japanese restaurants due to people's garbling and malicious dissemination of unverified remarks. They hope to get compensation through legal channels.

The third hot topic (as shown in Fig. 11) is different from the previous two stages, with core nodes such as Europe, Norway and conspiracy theory, which reflects the related issues summarized by netizens from various angles, such as "Beijing epidemic can be traced back to the European COVID-19 pedigree in March and April", "bad salmon farming environment in Norway" and "malicious poisoning by hostile forces". Except for the core nodes, the weight of each node in the figure is high, which highlights the profound influence of the topic to a certain extent and contains more high-value information.
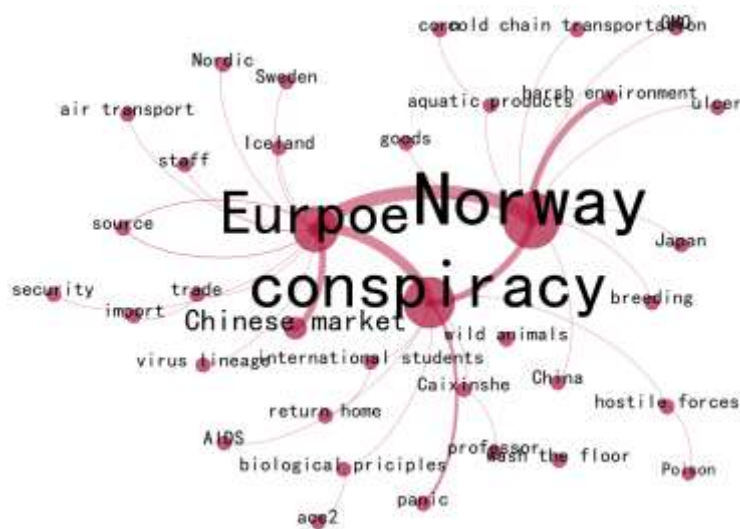

Fig 11: The third hot topic of peak stage

Fig 12: The third hot topic of dissipation stage

2.6.4 Hot topics in the dissipation stage

The popularity of the incident suddenly dropped since June 18. The core nodes of the first hot topic were salmon, hitting the shelf, innocence and rainbow trout, and the topics such as "salmon can be eaten", "salmon selling can be expected soon" and "domestic rainbow trout is delicious" were mainly discussed. The significant difference between this stage and the previous one is whether the topic of "whether the salmon can be eaten" has been re-mentioned, reflecting the netizens' review of the incident at the end of public opinion. The core nodes of the second hot topic are still vendors and Japanese restaurants, and the topic of when the salmon will be put on the shelves again was discussed. The third hot topic is shown in Fig. 12. The core nodes are the COVID-19 and inactivated vaccine, which indicates that netizens had gradually started to discuss the nature of the COVID-19 and the progress of vaccine development, sharply reflecting the subsequent impact of the event.

2.7 Analysis on the Evolution of Hot Topics in Different Stages of Public Opinion Incidents

According to the above analysis, the heat map of each stage of the public opinion event is obtained, and the hot core topics of each stage are summarized as shown in Table IV.

TABLE IV. Hot topics and statistics of all stages

Stage	Topic type	Node ratio	Node degree ratio	Topic summary	Topics covered
	One	42.36%	38.21%	virus found	Virus detected on salmon cutting board; Beijing-Shenzhen seafood market turmoil
infancy	Two	14.28%	18.38%	Check the seafood market	Salmon removed from Japanese food store; imported aquatic products checked in seafood market;
	Three	11.21%	5.21%	Import channel traceability	Salmon from Norway; Customs inspection of imported goods; Beijing

					traffic control
fermentation	One	45.32%	48.57%	virus infection mechanism	Is salmon a host; salmon has parasites; mammalian infection
	Two	20.78%	31.78%	grocery store loss	The Japanese food store suffered heavy losses; the Japanese food store sold rainbow trout; the price of salmon dropped by 70%
	Three	16.37%	8.24%	Foreign breeding environment	Salmon is imported from many European countries; Norwegian salmon farming environment is harsh; domestic supervision is unfavorable
peak	One	35.14%	37.44%	Can salmon be infected	The receptor of the new coronavirus is the ace2 protein; the possibility of salmon infection is almost 0
	Two	26.78%	16.21%	Event source interview error	Zhang Yuxi's interview was confusing; 40 positive samples only mentioned salmon; compensation for Japanese food store
	Three	14.28%	19.12%	Epidemic traceability	The epidemic in Beijing can be traced back to the European new crown virus lineage in March and April; the Norwegian salmon farming environment is harsh; hostile forces maliciously poisoned
dissipation	One	28.35%	29.26%	Salmon is back on the shelves	Salmon can be eaten; salmon is just around the corner;
	Two	18.21%	20.38%	Seafood market picks up	Japanese food stores put salmon on the shelves; the seafood market is picking up; salmon is not in the pot
	Three	15.47%	11.24%	Vaccine development progress	Clinical trial of new coronavirus inactivated vaccine; how long will the epidemic continue

The node number ratio in the table is the ratio of the total number of nodes in the topic to that in the current stage, and the node degree ratio is the ratio of the sum of the node degrees in the topic to that in the current stage, wherein the node degree refers to the total value of the external association weights of a node, and corresponding data is obtained by screening all relevant nodes and edges of different hot topics.

To some extent, the sum of node degrees reflects the popularity of the topic, which is used to reveal the evolution law of topics at all levels in this paper, and the relative ratio is adopted to eliminate the difference of data volume in different periods and the influence of clustering parameter changes on the results. Six representative hot topics were selected from each life cycle of public opinion for quantitative analysis. The sum of node degrees of hot topics was measured according to the sum of node degrees of core relevance and core expansion relevance corresponding to topics. The changing trend of hot topics in

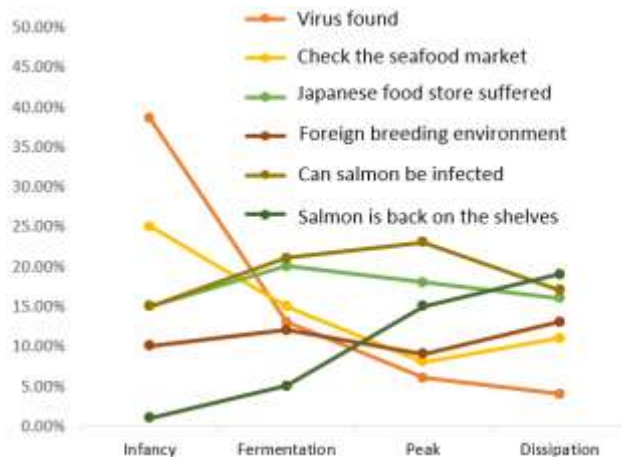different public opinion stages is shown in Fig. 13.

Fig 13: Evolution trend of hot topics

As shown in the figure, the node rate of "virus found" in the infancy stage was 38.5%, which was much higher than other topics, and then rapidly decreased, with the ranking and value in the dissipation stage reduced to the minimum. Similarly, the evolution law of "checking the seafood market" was similar to that of "discovering viruses", both of which caused intense discussion during the period as event sources and then rapidly declined. The change and ranking of the node degree sum of "Japanese restaurant loss" and "foreign aquaculture environment" in the four stages were relatively stable, with the average values of 17.3% and 11% respectively. The sum of node degrees of "whether salmon can be infected" had been steadily increasing, from the bottom at the infancy stage to the second place in dissipation stage. The popularity ratio of "salmon selling again" basically increased steadily in four stages, rising to the first place in dissipation stage.

III. CONCLUSION

In this paper, for the agricultural network public opinion events that have aroused widespread concern in all walks of life, by getting a large amount of relevant Weibo blog posts, dependency parsing rules were used in the short text analysis of public opinion information, so that the extracted hot words have a clear semantic relationship. At the same time, the number of likes and forwards of published information were included in the calculation of correlation weights between words, and social network method was used to realize the discovery and visual analysis of public opinion hot topics. According to the evolution of hot words and their network topology in the process of public opinion development, the reasonable division of public opinion life cycle and the accurate refinement of public opinion topics were realized, and the evolution law of public opinion hot topics was quantitatively analyzed by combining the characteristics of the number and degree of nodes of social networks. Based on the data sample of 17,579 blog posts under the Salmon COVID-19 incident, the relationship network diagram with clear and prominent themes was obtained, thus the hot topics in each life cycle of public opinion can be quickly analyzed, which is in line

with the trend of real public opinion events. The findings of the research are

(1) The causes of public opinion events only aroused heated discussion at the initial stage, and were reviewed at the end of public opinion, while the popularity remained low in the rest stages.

(2) There is no obvious fluctuation in the popularity of hot topics related to all key facts.

(3) External stimuli had a strong impact on the evolution of public opinion. "COVID-19 receptor is ace2 protein, and salmon can also be infected with virus" had significantly increased the popularity of the topic at the individual level, thus inhibiting its popularity at the organizational and social levels.

(4) Topics at the organizational level are more likely to be discussed in small groups at the infancy stage of public opinion, but the popularity will continue to rise in the future due to its great connection with the nature of events.

In this paper, the method of combining visual analysis with quantitative analysis, which is based on the experimental data of mainstream media texts during the response to the COVID-19 epidemic, has good universality, and can provide reference for the research of unexpected public opinion events in other fields in the future. However, this paper has certain limitations: although the experiment is mainly data-driven, there are some subjectivity in the selection of the number of topics and the selection and judgment of topics, which need to be further explored in the future related research.

ACKNOWLEDGEMENTS

REFERENCES

[1] Ma N, Liu Y, Lian Y (2019) A review of public opinion risk for unexpected emergency. Journal of Intelligence 38(06):88-94.

[2] Yi X. Research on college news topics discovery based on LDA topic model. Beijing Jiao tong University, Beijing, 2019

[3] Zhou K. Weibo hot topic discovery and analysis based on LDA topic model. Southeast University, Nanjing, 2018

[4] Liu H. Research on hot topic detection based on model fusion. Shandong University of Science and Technology, Shandong, 2019

[5] Xiao Q, Xie H, Liu P (2019) A recognition method of hot events in social media by integrating LDA into CNN. Information Science 37(11):27-33

[6] Zheng H, Liao C, Li T (2019) A topic detection method for network long text. Chinese Journal of Engineering 41(09):1208-1214

[7] Ding S, Wang P, Gong S (2018) Research on potential subject discovery of network public opinion based on community discovery and keyword co-occurrence — Sina micro-blog Wei Zexi incident as an example.

Information Science 2018, 36(07):78-84

[8] Li L, Liu J, Zhang H (2016) Topics identification and evolution trend of network public opinion based on co-occurrence analysis. Information Science 34(01):44-47+57

[9] Yuan J, Du B (2021) Analysis of hotspots in domestic supply chain research field based on co-word analysis. Logistics Technology 40(02):96-101

[10] Bi L, Wang T, Zuo W (2014) Empirical study of micro-blog hot topic identification based on probability model. Information Studies: Theory & Application 37(02):112-116

[11]An L, Wu L (2017) An integrated analysis of topical and emotional evolution of microblog public opinions on public emergencies. Library and Information Service 61(15):120-129

[12] Liu W. Representation learning and dependency syntax for text summarization. Shandong Normal University, Shandong, 2020

[13] Liu H. Chinese dependency parsing based on deep learning. Beijing Jiao tong University, Beijing, 2019

[14] An L, Ou M (2017) Social Network Sentiment Map of the Stakeholders in Public Health Emergencies. Library and Information Service 61(20):120-130

[15] Liang X, Xu J (2018) Sentiment Analysis of Objects in Public Opinion Events and Their Relation Network Research. Information Science 36(02):37-42.

[16] Teng L, Shen J, Gao J (2013) Cognitive Structure in co-word knowledge network: theory, method and a case study. Journal of the China Society for Scientific and Technical Information 32(09):976-989

[17] Zhuo M, Kong X (2016) Wages of Migrant Workers and Social Networks - The Cause Effect Analysis Based on the Relationship Strength and Relational Resources. Economic Survey 33(06):48-53

[18] Chu D. Cluster analysis and its applied research. Jiangnan University, Wuxi, 2020

[19] Han K, Xing Z, Liu Z, et al (2021) Research on Public Opinion Analysis Methods in Major Public Health Events: Take COVID-19 Epidemic as an Example. Journal of Geo-information Science 23(02):331-340

[20] Wang J, Xia Z, Wang Y, et al (2020) Visualization and Evolution of hot topics of internet public opinion based on syntax rules and social network analysis. Information Science 38(07):132-139