

Analysis of News Dissemination Characteristics under the Trend of Short Video Convergence Based on Multi-Priority Dynamic Threshold Algorithm

Huan Xie^a

^a*School of Humanities and Media, Pingxiang University, Pingxiang, Jiangxi, 337055, China*

Abstract. The rapid development of science and technology and the continuous iteration of mobile terminals have led to great changes in the news dissemination mode. After summarizing the characteristics of news publishing in the short video perspective, the study constructs a behavioral communication characteristics analysis model based on multi-priority dynamic thresholding. The model is able to automatically cluster according to the characteristics of the dataset by establishing a dynamic threshold selection model, which eliminates the difficulty of setting parameters manually. Meanwhile, the text similarity calculation based on the kernel method is used to realize the mining of short video news information. Conducting experimental analysis and practical application of the model's performance, it is found that the scores of the multi-priority dynamic threshold model on AR, PR, RE and F-score are 0.71, 0.67, 0.72 and 0.68, which are better than those of the comparison model, reflecting the model's good ability to mention information of short news videos. Who (97.21%) and What (100%) elements are the highest in news short video titles, and the proportion of monosyllabic and disyllabic words is 93.2%. It was found that short video news is characterized by fragmentation, concise content and interactivity.

1. Introduction

With the iterative upgrading of new media technology, cutting-edge technologies such as VR, sensors, artificial intelligence and other cutting-edge technologies continue to flood into the public eye, subverting the traditional way of news production and reconfiguring the shape of news business. Traditional media have strived for breakthroughs, exploring the road of breakthrough and accelerating transformation and upgrading in the new communication ecology [1–3]. Under the perspective of media convergence, the traditional graphic reports just do not have the competitiveness and eye-catching power, short video with strong social, interactive, fragmented, suitable for mobile scene dissemination advantages, the rapid rise of the wave of momentum, become the main way of news release [4, 5]. In the context of the overall user scale of the mobile Internet continues to narrow the increase, the user scale and viewing time of short video continues to grow. Short video, a new form of media, has been increasingly emphasized by the

Corresponding author: Huan Xie, mail address: 13672221934@163.com

Received: 15 November 2024; Accepted: 13 January 2025; Published: 12 March 2025

Keywords. Multiple prioritization; Dynamic thresholding; Text similarity; Information extraction; Short video news

2010 *Mathematics Subject Classification.* –

Cited this article as: Xie. H, Analysis of news dissemination characteristics under the trend of short video convergence based on multi-priority dynamic threshold algorithm. *Engineering and Its Applications*, 2(1), 2025, 1-22.

© 2026 Engineering Journals. Published by Engineering Journals

news industry, and many traditional media have launched short video news boards, online short video news clients, and produced short video news columns, which has set off a boom in the development of short video news [6–8]. Under the trend of short video convergence, short video news will undoubtedly bring more possibilities for the information dissemination ecology, provide users with more optimized service experience, bring a new round of iterative upgrading for the industry, and create unprecedented communication value for the media ecology [9, 10].

In the field of news communication, in the face of the increasingly changing communication pattern, innovation is the key to improve the core competitiveness of news media. Literature [11] takes media convergence as an entry point to explore the impact of the development of new media on traditional media, and puts forward a series of adjustment strategies to make traditional media invincible in the fierce competition, and continue to provide the public with high-quality and reliable information services. Literature [12] pointed out that under the drive of scientific and technological development, traditional media must follow the trend of the times to carry out transformation and upgrading in order to be invincible, and the successful practice of “Kanghui VLOG” undoubtedly provides a reference and inspiration for other counterparts. Only by constantly innovating, broadening the means of reporting and content areas, and effectively integrating them can we win the support of the audience and remain competitive. Literature [13] uses AISAS theoretical model and K-Means optimization algorithm with high efficiency and accuracy to analyze the effectiveness of short video news dissemination, which provides a feasible solution for further mining and utilizing the massive information in the field of short video news. Literature [14] introduced a multi-priority dynamic threshold algorithm in the transmission of short video news, and verified through experiments that the short video news mode has more prominent advantages than the traditional news dissemination mode, and that the short video news mode can satisfy the user’s demand for instant information acquisition and improve the efficiency of information dissemination, which shows the prospect of wide application. Literature [15] applies the intelligent neural network in artificial intelligence (AI) technology to intelligent robots or other related systems to solve the problem of information acquisition and dissemination efficiency in the field of short-video news, so as to promote the development of short-video news network.

This paper summarizes the communication characteristics of the current short video news in combination with the trend of short video convergence. By constructing a dynamic threshold selection method for automatic selection of algorithm parameters, improving dynamic nearest neighbor selection and automatic selection of relative density on the basis of RDBKNN algorithm, and selecting the kernel method as the calculation method of text similarity, we propose a behavioral propagation characteristics analysis model based on multi-priority dynamic thresholding. The analysis experiment of dynamic threshold is implemented, and at the same time, the multi-priority dynamic threshold model of this paper and other models are compared and analyzed for multiple evaluation indexes of short video news information extraction, to verify the short video news information mining effect of the model of this paper. Finally, the short video content of a news is collected as a research sample, and the model is used to conduct lexical analysis, word frequency analysis and word length analysis of news headlines at a time to analyze the characteristics of news dissemination under the convergence of short videos.

2. Characteristics of news dissemination in the view of short video

With its unique form and rapid dissemination, short video has changed the mode of traditional news reporting, and at the same time brought more innovative possibilities. Based on this, this chapter discusses the characteristics of news dissemination in the view of short video.

2.1. Simple and intuitive

The simple and intuitive way of communication is a major feature of short videos in news dissemination. The length of short videos is usually between tens of seconds and a few minutes, which is easier to attract the attention of viewers than traditional text-based news reports. When using short videos for news dissemination, producers need to carefully plan and edit the content to ensure that the main points of

the news can be accurately and completely conveyed within a limited time. In addition, short videos can enhance the vividness and comprehensibility of the content by employing a variety of visual effects, such as exciting graphics, eye-catching charts and diagrams, exciting sound effects, as well as clear and concise subtitles and copy.

2.2. Content diversity

Short videos make full use of multimedia technology, integrating text, pictures, video, audio and other forms together to provide viewers with a rich and diverse content experience. Watching short videos can make the audience feel the authenticity and urgency of the event, and this intuitive way of communication can enhance the audience's sense of trust and participation in the news. In addition, the diversified content form of short video is not limited to the use of multimedia, but also includes the application of a variety of creative techniques. For example, the producer can use animation, cartoon, simulation scenes and other elements to show the audience news events or information in a unique and interesting way. Through this creative form of expression, short videos can not only convey information, but also bring the audience more artistic and entertaining experience.

2.3. Highly interactive

Through the interactive function of the short video platform, viewers can comment, like, share and other operations to communicate and interact with the producer and other viewers. This interactivity can greatly enhance the audience's participation and attention to the news content. Viewers can express their views on news events, share their experiences or ask questions through comments. By communicating and discussing with other viewers, viewers can understand the meaning behind news events more comprehensively and form more reasonable judgments and opinions. At the same time, the audience's behavior such as liking and sharing can also help spread the news content and promote the dissemination of information and influence. In order to further promote user participation, some short video platforms also carry out activities such as voting surveys and virtual gift giving.

3. Multi-priority dynamic threshold-based news dissemination characterization model

Under the trend of short video convergence, this chapter proposes a multi-priority dynamic threshold-based model for news dissemination feature analysis. The algorithm first establishes a dynamic threshold selection model with hierarchical threshold selection for automatic selection of algorithm parameters. Secondly, it is improved on the basis of the RDBKNN algorithm, using the dynamic threshold selection model to select the respective appropriate nearest neighbors for each data point to form a more natural neighborhood. The selection of relative density threshold parameters is also selected using the dynamic threshold selection model, but a different strategy is adopted for execution.

3.1. Dynamic threshold selection method

The threshold selection method in the dynamic threshold selection model is as follows:

1. First, the data set S is divided into n subsets with the following division rules:

$$\forall q_i \in S, \quad S_1 = \{q_i \mid q_i < \theta_1\}, \quad S_k = \{q_i \mid \theta_k > q_i \geq \theta_{k-1}\}, \quad S_n = \{q_i \mid q_i \geq \theta_{n-1}\} \quad (1)$$

where $\theta_k = k\delta/n$, $k = 2, \dots, n-1$, $\delta = \max(S) - \min(S)$.

The number of subsets n is determined using the following method: let the average number of samples in a single interval be avg_n and the total number of samples be N , then, as N increases, avg_n must be a non-degenerate positive integer degree series satisfying the following 3 conditions:

- A. $\lim_{N \rightarrow \infty} avg_n = \infty$.
- B. $\lim_{N \rightarrow \infty} avg_n/N = 0$.

- C. $\lim_{N \rightarrow \infty} avg_n / \log(N) = \infty$.

The functional relation $avg_n \propto \sqrt{N}$ is a special solution satisfying the above condition, so it can be taken as $avg_n = \sqrt{N}$.

From this, the number of subsets n is then divided for S using an exhaustive search method that gradually increases from 2 until any one of them, subset S_i , terminates when condition $|S_i| \leq N/n$ is satisfied. The value of n obtained at this point is the number of subsets.

2. Define a variable scale J as the demarcation point, and take different processing strategies for subsets S_k and S_m of $k > J$ and $m \leq J$, respectively, which are used to regulate the roles of high-value data and low-value data.

3. The data points contained in subset S_n have high values, and it can be assumed that most of the data points in S_n are data that satisfy the requirements, so the threshold value should be lowered so as not to remove too many valid data points. The threshold value is selected as:

$$t_n = \theta_{n-1} + \sigma_n \sqrt{2 \log(|S_n|) / |S_n|} \quad (2)$$

where σ_n is the standard deviation of data points in S_n and $|S_n|$ is the number of data points in set S_n .

4. Some of the data points in subset S_k ($k = J + 1, \dots, n - 1$) have high values, while others may be low, and the threshold values should be increased appropriately to retain as many high value data points as possible. The selected threshold value is:

$$t_k = \theta_{k-1} + \sigma_k \sqrt{2 \log(|S_k|) \ln(k + 1)} \quad (3)$$

where σ_k is the variance of the data points in S_k and $|S_k|$ is the number of data points in set S_k .

5. Subset S_m ($m = 1, 2, \dots, J$) contains data points with very different values compared to other subsets, and generally contains only very low values, so it is treated differently from sets with subscripts greater than J :

- A. Firstly, merge all the subsets that satisfy condition $m \leq J$ to construct a new subset S' .
- B. Increase the threshold to remove as many of the low value data points as possible. The selected threshold value is:

$$t_m = \text{mean}(S'_j) + \sigma'_j \sqrt{2 \log(|S'_j|)} \quad (4)$$

where $\text{mean}(S'_j)$ is the mean of subset S'_j , σ'_j is the variance of data points in S'_j , and $|S'_j|$ is the number of data points in set S'_j .

6. The mean of the set of data points selected in the above steps is taken as the final threshold value thr . The formula is:

$$thr = \frac{\sum_{p_i \in S'} p_i}{|S'|} \quad (5)$$

where $|S'|$ is the number of data points in set S' .

3.2. Parameter-free local density clustering algorithm

3.2.1. Dynamic nearest neighbor selection

In the parameter-free local density clustering (DTSLD) algorithm, the global parameter k in the RD-BKNN algorithm is discarded, and instead the method of dynamically selecting an appropriate number of neighbors for each data point is used to construct the neighborhood. Since the nearest neighbor selection should try to exclude neighbors with low similarity, for this consideration, the scale J is set as the subscript of the subset with the largest variance in the dynamic threshold selection model.

The methods of dynamic nearest neighbor selection for each data point are summarized as follows:

1. Without loss of generality, set the current data point as p .
2. Select all neighbors with similarity greater than 0 to form an initial set of neighbors S .
3. Threshold the set S using the dynamic threshold selection model to obtain the threshold θ : Scale J is set to:

$$J = \arg \max(\sigma_j), \quad 1 \leq j < n \quad (6)$$

where n is the number of subsets divided in the dynamic threshold selection model and σ_j is the variance of subset S_j .

4. Select all the data points among them that satisfy condition (7) as the final set of neighbors.

$$\text{sim}(q_i, p) \geq \theta \quad (7)$$

3.2.2. Automatic selection of relative density thresholds

After constructing the neighborhood for each data point, the calculation of neighborhood density and relative density can be completed, and then the core object can be selected. In this paper, the formula for calculating relative density is modified as follows.

$$rd(p) = \sum_{q \in N\text{Neighborhood}(p)} \frac{\min(\text{density}(q), \text{density}(p))}{\max(\text{density}(q), \text{density}(p))} \quad (8)$$

where $\text{density}(p)$, $\text{density}(q_i)$ are the neighborhood densities of data points p , q_i , respectively, and $N\text{Neighborhood}(p)$ is the set of neighbors of p .

After this improvement, the value range of $rd(p)$ is $(0, 1]$, which still satisfies the condition that the relative density between neighbors with close density is close to 1, and meets the condition of dynamic threshold selection.

In this paper, the scale J is set to 1, balancing the role of high-value data points and low-value data points, and the dynamic threshold selection model decides the threshold value by itself according to the characteristics of relative density changes.

The method of relative density threshold selection is summarized as follows:

1. Construct set S from the relative density of all data points $rd(p_i)$.
2. The scale J is set to 1, and the threshold value δ is obtained by using the dynamic threshold selection model to select the threshold value for the set S .

Since the selected relative density threshold is a value close to 1, which is different from the threshold in the original RDBKNN algorithm, the conditions for the selection of the core object are changed as follows:

$$rd(p) \geq \delta \quad (9)$$

3.2.3. Basic definition of DTSLD algorithm

The above two sections illustrate the improvements of the DTSLD algorithm over RDBKNN, and this section gives the complete definition of the DTSLD clustering algorithm. Given dataset D , there are:

- 1. Neighborhood.** The neighborhood of data objects $p \in D$, p is denoted as $N\text{Neighborhood}(p)$, and the data objects belonging to $N\text{Neighborhood}(p)$ satisfy the following conditions:

$$N\text{Neighborhood}(p) = \{q_i \in D / \{p\} \mid \text{sim}(p, q_i) \geq \text{thr}(p)\} \quad (10)$$

where $\text{sim}(p, q_i)$ is the similarity between p and q_i , and $\text{thr}(p)$ is determined by the dynamic threshold selection model.

- 2. Neighborhood density.** For data objects $p \in D$, p , the domain density is denoted as $\text{Density}(p)$ and defined as:

$$density(p) = \frac{\sum_{q_i \in NNeighborood(p)} sim(p, q_i)}{|NNeighborood(p)|} \quad (11)$$

where $|NNeighborood(p)|$ is the number of neighbors in the neighborhood of p . The domain density of data object p indicates the mean value of similarity between the data objects in the neighborhood point set of p and p .

3. The relative density of data object p about its neighbors $NNeighborood(p)$. The relative density of p about its neighbors is denoted as $rd(p)$ and is defined as follows:

$$rd(p) = \sum_{q \in NNeighborood(p)} \frac{\min(density(q), density(p))}{\max(density(q), density(p))} \quad (12)$$

$rd(p)$ reflects the difference between the near-neighbor density of data object p and the near-neighbor densities of its neighbors, and when $rd(p)$ is close to 1, it indicates that there is a small difference in density between data object p and its neighbors, and that they are very close to each other in terms of the density of the data distribution.

4. Core object. Data object p is referred to as the core object and should satisfy the condition: $rd(p) \geq \delta$, where $\delta > 0$. δ is the relative density threshold, which is set to be selected automatically using the dynamic threshold selection model.

5. Relative density of data object p with respect to class C . If data object p is the core object, the relative density of p about class C is noted as $rd_c(p)$ and defined as follows:

$$rd_c(p) = \frac{\min(density(q), density(p))}{\max(density(q), density(p))} \quad (13)$$

Of these, $density(p) = \max_{p_i \in C} (density(p_i))$.

6. Direct near-neighbor reachability. For data objects $p, q \in D$, data object q is directly nearest neighbor reachable from object p if $q \in NNeighborood(p)$, and p is a core object.

7. Near-neighbor connectivity. If there exists a chain of data objects: $p_1, p_2, \dots, p_n$ ($p_1 = p$, $p_n = q$), for $p_i \in D$ ($1 \leq i \leq n$, and p_i is a core point), p_{i+1} is directly near-neighbor reachable from p_i , and both satisfy the condition: $rd_c(p) \geq \delta$, where $\delta > 0$. Then object q is near-neighbor connected from object p .

3.3. Text similarity calculation based on kernel methods

DTSLD clustering algorithm is applied to the news text data for topic structure mining, the problem that needs to be solved is the calculation method of text data similarity. In this paper, the kernel measure is used to calculate the similarity of text data, and the hybrid kernel function construction idea is shown in Figure 1, which selects the improved latent semantic kernel (LSK) and the composite kernel ANOVA for weighting to construct a new kernel function for calculating the text similarity.

3.3.1. Constructing an improved latent semantic kernel

TF-IDF is a commonly used document vectorization method in VSM modeling, which integrates the frequency of occurrence of a word in a single document and the frequency of occurrence of the word in a collection of documents to model the text as a word frequency vector. In this paper, text vectors are represented based on this method.

The entire set of questions q and the retrieval results d_i ($i = 1, \dots, M$, a total of M retrieval results) of each dataset is called a corpus, and the set of words appearing in the corpus is called a lexicon, and given a retrieval result, which is considered as a bag of words, it is denoted as a row vector, each of whose dimensions is associated with one of the words of the lexicon. The retrieval result d and question q are represented by mapping:

$$\phi : d \rightarrow \phi(d) = (tf(t_1, d), tf(t_2, d), \dots, tf(t_N, d)) \in \mathbb{R}^N \quad (14)$$

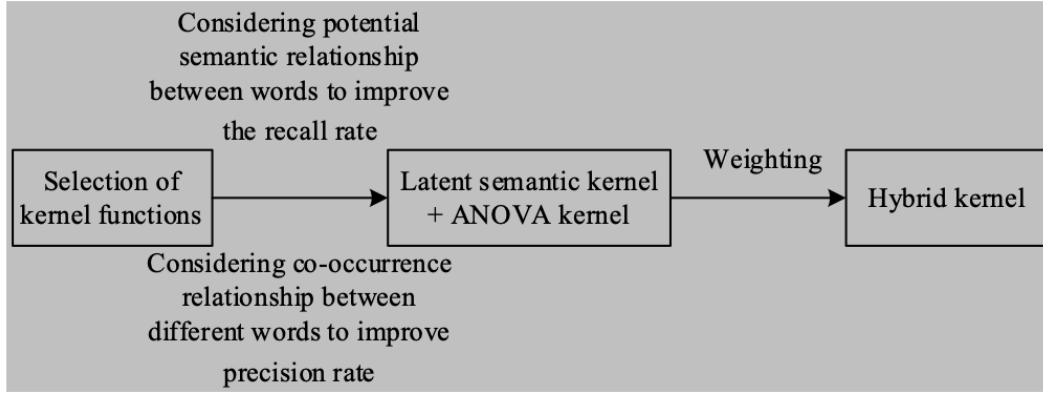


Figure 1: The idea of a hybrid nuclear function

$$\phi : q \rightarrow \phi(q) = (tf(t_1, q), tf(t_2, q), \dots, tf(t_N, q)) \in \mathbb{R}^N \quad (15)$$

where $tf(t_i, d)$ is the frequency of occurrence of word t_i in document d , and $tf(t_i, q)$ denotes the frequency of occurrence of word t_i in question q , usually 1. The size of the lexicon N is usually a very large number. The matrix of the word document for an anthology with l documents as row indicators and words of the anthology as column indicators is denoted as:

$$D = \begin{bmatrix} tf(t_1, d_1) & \cdots & tf(t_N, d_1) \\ \vdots & \ddots & \vdots \\ tf(t_1, d_l) & \cdots & tf(t_N, d_l) \end{bmatrix} \quad (16)$$

Then its kernel matrix is $K = DD'$ and its kernel function is

$$k(q, d) = \langle \phi(q), \phi(d) \rangle = \sum_{j=1}^N tf(t_j, q) tf(t_j, d) \quad (17)$$

However, the bag-of-words representation does not take into account the semantic information of words. In order to solve this problem semantic kernels need to be constructed on the basis of bag-of-words representation. Different words have different levels of importance to the topic. The frequency of occurrence of a word in a document is used to quantify the importance of the information carried by the word, i.e., the IDF rule, which is expressed as:

$$w(t) = \ln \left(\frac{l}{df(t)} \right) \quad (18)$$

Where l is the number of documents present in the anthology, $df(t)$ is the number of documents containing the word t , and $w(t)$ is the absolute scale measuring the weight of the word t as defined by the inverse document frequency IDF rule. The matrix R is chosen to be the diagonal matrix whose elements are $R_{tt} = w(t)$, and thus the mapping becomes $d \rightarrow \varphi(d) = \phi(d)R$. The kernel with the information about the weights of the words is:

$$k_0(d_1, d_2) = \phi(d_1)RR'\phi(d_2)' = \sum_t w(t)^2 tf(t, d_1)tf(t, d_2) \quad (19)$$

On the other hand, there may be potential semantic relationships between different words. Considering that there is still a possibility of correlation between two question documents without a common word, i.e., there is a semantic similarity between different words, when two words i and j appear simultaneously in a

document is associated with some kind of relationship strength given by the frequency of co-occurrence of the words and the number of times of co-occurrence. When there is at least one document in the anthology in which the i th word occurs at the same time as the j th word by:

$$K_{ij} = (D'D)_{ij} = \sum_d tf(i, d) tf(j, d) \quad (20)$$

It is known that the (i, j) element of the word $\times$ word matrix $D'D$ is non-zero.

The number of words in the above dictionary is a very large value, and the number of documents in the retrieval results is often also a huge value. In order to reduce the overhead of computer computation and improve efficiency, come to use kernel principal component analysis. Take the kernel matrix $D'D$ of the first m eigenvalues corresponding to the eigenvectors (normalized) expressed as a column matrix U_m , the document will be projected to U_m of the column vector tensor into the space, the potential semantic kernel:

$$\phi(d)U_m = \lambda_i^{-1/2} \sum_{j=1}^l (v_i)_k (d_j, d) \Big|_{i=1}^m \quad (21)$$

Where λ_i, v_i are the i th eigenvalue and eigenvector pairs corresponding to kernel matrix $D'D$. Through this principal component analysis, the space tensored by the first m eigenvectors is taken as the mapping space to achieve the purpose of dimensionality reduction and to reduce the time and null direction overhead of the computation.

Combined with the weight information representation of words, the new improved potential semantic kernel function is obtained as:

$$\phi(d)RU_m = \lambda_i^{-1/2} \sum_{j=1}^t (v_i) \sum_1 w(t)^2 tf(t, d_j) tf(t, d) \Big|_{i=1}^m \quad (22)$$

Then the degree of similarity between the calculated search result d and question q can be expressed as:

$$k(q, d) = \left\langle \left(\lambda_1^{-1/2} \sum_{j=1}^l (v_i), \sum_1 w(t)^2 tf(t, q_j) tf(t, q) \right)_{i=1}^m, \left(\lambda_1^{-1/2} \sum_{j=1}^l (v_i), \sum_1 w(t)^2 tf(t, d_j) tf(t, d) \right)_{i=1}^m \right\rangle \quad (23)$$

Remove the effect of differences in text length on the results by unitizing them:

$$\hat{k}(q, d) = \frac{k(q, d)}{\sqrt{k(q, q) k(d, d)}} \quad (24)$$

3.3.2. Selection of ANOVA composite kernel

A more-concordant ANOVA kernel is used to express the effect of different question words co-occurring with the degree of relevance of their retrieval results to the question, which is briefly described as follows. The original features in the input space are combined and their combined features can be defined by the following equation:

$$\phi(x) = \left\{ \phi_I(x) = x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n} \mid I = (i_1, i_2, \dots, i_n) \in \text{CondSet} \right\} \quad (25)$$

Where $I = (i_1, i_2, \dots, i_n)$ denotes the power of each component in the feature combination, and $I \in \text{CondSet}$ denotes that the vector has to satisfy certain constraints. For a certain combination of features $\phi_I(x)$, if any of the original features x_i is 0 (i.e., the search question word does not appear), the entire combination of features does not exist, so that the retrieval results (document d) are not involved in the calculation. According to the kernel function theory, given a feature space there must be a kernel function corresponding to it. So the kernel with the ability to generate the above features is selected.

The feature set in the feature space determined by the r th order ANOVA kernel consists of all monomials of the r th order to the 1st power, and the definition of the kernel function and its corresponding mapping are shown in the following equation:

$$k_A^{(r)}(x, z) = \sum_{1 \leq j_1 < j_2 < \dots < j_r \leq s} \left(\prod_{k=1}^r x_{j_k} z_{j_k} \right) = \left\{ \phi_I(x) = x_1^{i_1} x_2^{i_2} \dots x_n^{i_n} \mid I = (i_1, \dots, i_n) \in \{0, 1\}^n, \sum_{j=0}^n i_j = r \right\} \quad (26)$$

In some cases the search terms do not need to be co-occurring, and their retrieval results have some similarity to the questions asked, so the standard ANOVA kernel of first order $K_n^1(x, z)$ was considered at the same time.

where $K_n^s(x, z)$ denotes the standard ANOVA kernel of order r for the original space with n input features: $\sum_{i=2}^r K_n^s(x, z)$ denotes the combination of features of each order (not higher than order r). At the same time, the new kernel $K_n^s(x, z)$ is constructed by transforming the standard ANOVA kernel of each order through the operator Trans to generate a new nonlinear space, which overcomes the problem of nonlinear relationships between feature combinations in the feature set generated by the standard ANOVA kernel.

3.3.3. Constructing new kernel functions

Kernel functions are closed: if k_1, k_2 are kernel functions, then $a_1 k_1 + a_2 k_2$ is also a kernel function if and only if $a_1 \geq 0$, and $a_2 \geq 0$. Based on this property of kernel functions, existing kernel functions can be utilized to construct new composite (or hybrid) kernel functions. Using this construction technique of kernel function, the composite kernel is constructed by simultaneously considering the potential semantic information between different words, the role of question word co-occurrence on relevance, and the significance of the occurrence of a single question word on relevance:

$$K_c(x, z) = \delta_1 \hat{k}(x, z) + \delta_2 \text{Trans} \sum_{p=2}^n K_n^3(x, z) + (1 - \delta_1 - \delta_2) K_{n1}(x, z) \quad (27)$$

where δ_1, δ_2 ($0 < \delta_1 < 1$, $0 < \delta_2 < 1$) are the weighting factors.

3.4. Experiments and analysis of results

3.4.1. Data sources

In this paper, 120 short news videos are collected from online platforms to form a dataset, which is classified into a collection of textual frames and a collection of non-textual frames using OCR text recognition. Users manually create video summaries and automatically generated for comparison by 50 staff members manually create video summaries for different videos, each video has five different manually produced summaries, in order to facilitate the subsequent experiments to compare the five different summaries, through the above method for filtering and merging, the merger results through the manual verification.

3.4.2. Dynamic threshold analysis experiments

In this experiment, the effect of text similarity threshold on the quality of news short video information extracted by the model is investigated, and the value range of text similarity threshold is set to 0.6-1.0 in this paper, and selected with a step size of 0.1. The news information extraction under different thresholds is shown in Figure 2, which shows the Precision, Recall and F-score of the model under different thresholds, respectively. As the threshold value increases, the precision of the model for news information extraction decreases, the values of Recall and F-score increase, and the rising trend of F-score and the decreasing trend of precision slow down when the threshold value is greater than 0.9. For short video news, the model in this paper can control the precision, recall, and F-score values at a high level when the threshold is 0.9.

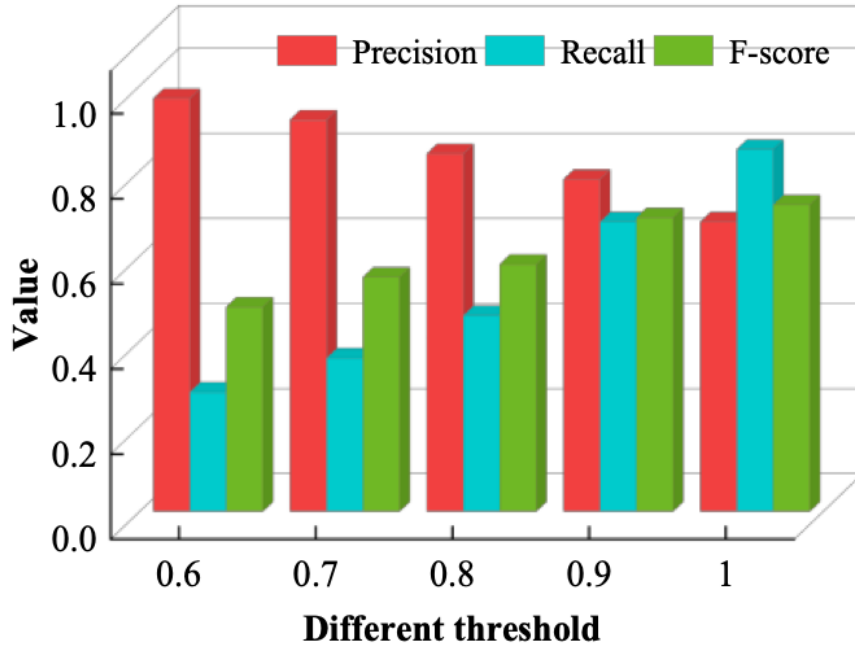


Figure 2: Information on information extraction under different thresholds

3.4.3. Comparative analysis of models

This paper proposes a multi-priority dynamic thresholding model to analyze the characteristics of news dissemination in the perspective of short videos, and in order to prove the effectiveness of this paper's algorithm, the model is compared with two more classical clustering-based video extraction algorithms, VSUMM and DT. Table 1 shows the detailed results of various methods on the benchmark dataset, including the evaluation results of accuracy (AR), error rate (ER), precision (PR), recall (RE), and F-score.

The scores of this paper's multi-priority dynamic thresholding model on AR, PR, RE, and F-score are better than the other methods, with the score values of 0.71, 0.67, 0.72, and 0.68, respectively. This indicates that the short video news text extracted by this paper's method contains more effective information. Meanwhile, this paper's multi-priority dynamic thresholding model makes a good trade-off between AR and ER, which means that it not only accurately describes the video content, but also generates less redundancy. In addition, the best F-score value obtained by this paper's model on the benchmark dataset also verifies that this paper's method improves the extraction accuracy of short video news information compared with several other algorithms.

Table 1: Detailed results of various methods in the benchmark data set

Models	AR	ER	PR	RE	F-score
DT	0.23	0.15	0.59	0.21	0.35
VSUMM	0.24	0.12	0.41	0.27	0.47
PHA	0.37	0.27	0.64	0.36	0.53
FTC	0.59	0.43	0.61	0.55	0.62
Ours	0.71	0.32	0.67	0.72	0.68

4. Analysis of the characteristics of news dissemination under the convergence of short video

The chapter takes the short video content of an official news as the research object, uses the multi-priority dynamic threshold model for information extraction of short video news content, and analyzes the characteristics of news dissemination under the trend of short video convergence.

4.1. Lexical analysis of news headlines

Almost all people devoted to journalism research or news report writing follow the doctrine of 5Ws (Who, What, When, Where, Why) elements of news as the basic guideline of news structure to dig out the specific meanings of news topics from different perspectives. Based on the model analysis of 120 short news videos, it is found that 97.21%, 100%, 65.47%, 62.62% and 34.52% of the news contain Who, What, Where, When and Why elements, respectively, and the four elements of Who, What, Where and When have a greater weight in comparison with Why. Where, When and Why elements in short video news is not particularly high, that is, most of the short video news only contains 1~3 of the five elements, reflecting the current short video news content of concise and fragmented characteristics. The statistical results of the lexical analysis of news headlines are shown in Table 2. From the table data, it is not difficult to find that the core four elements of the news are mainly concentrated from five types of words, i.e., nouns (905), verbs (524), non-predicate adjectives (42), tense words (95), and number words (16).

Table 2: Statistical results of the lexical analysis of news titles

Elements	Noun	Verb	Adjective	Tenses	Numerals	Other	Total
Who	563	0	30	3	5	11	612
What	0	524	0	0	0	37	561
Where	342	0	12	0	0	0	354
When	0	0	0	92	11	0	103
Total	905	524	42	95	16	48	

4.2. Word frequency analysis

Word frequency is the basis of linguistic quantitative research, including the calculation of frequency and frequency. Frequency refers to the number of occurrences of the subject within the corpus of the investigation, which is a specific number, which can intuitively reflect the real and original use of a word in the corpus.

Table 3 shows the statistics of high-frequency words in the sample short video news. Among the top 16 high-frequency words, there are both real words and imaginary words, with the highest frequency being the imaginary word “of”, which appears 641 times, followed by the real word “day”, which appears 452 times. In addition, a large number of words such as “day”, “month”, “year”, and “hour” appear due to the official and accurate characteristics of the content published by the sample news short video, and the time and place information need to be indicated.

4.3. Word length analysis

The word length distribution of short video news is shown in Figure 3. The share of mono- and disyllabic words in the corpus is as high as 93%, far exceeding that of polysyllabic words. Words of four syllables and above have extremely small values in terms of variety, number and cumulative coverage, and their usage is much less flexible than mono-, disyllabic and trisyllabic words. Monosyllabic and disyllabic words mostly appear in short video news, reflecting the short and interactive nature of short video news. Short video platforms have rich interactive features that can promote interaction between audiences and content creators. The audience can express their views, emotions and attitudes through comments, likes, shares and other functions.

Table 3: The high frequency word statistics of sample myopia news

NO.	Words	Frequency	NO.	Words	Frequency
1	The	641	9	Not	356
2	Day	452	10	You	338
3	At	402	11	Have	294
4	One	395	12	People	283
5	Month	387	13	Will	239
6	Is	368	14	This	222
7	Both	367	15	Know	220
8	And	362	16	Time	207

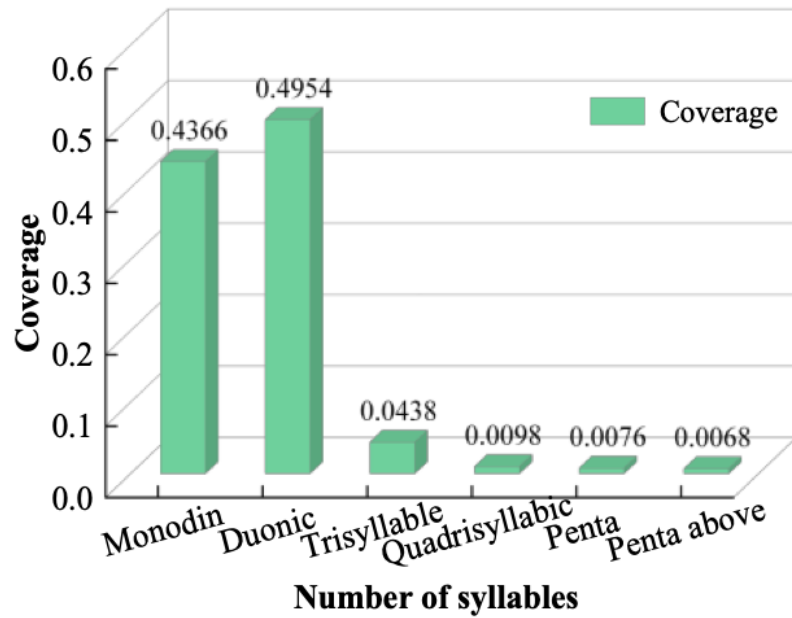


Figure 3: Long term distribution of short-frequency news

5. Conclusion

The emergence of short videos has injected new vitality into the production and production of news, realizing the innovation of news dissemination form and dissemination content. This project proposes a multi-priority dynamic thresholding-based behavioral communication feature analysis model, verifies its performance in short video news information extraction through experiments and model comparisons, and dissects the characteristics of short video news.

The multi-priority dynamic thresholding model in this paper has a better short video news information extraction effect, with scores of 0.71, 0.67, 0.72 and 0.68 on AR, PR, RE and F-score, and its performance is better than the comparison algorithm in terms of accuracy and precision.

Among the sample short video news, Who (97.21%) and What (100%) accounted for the highest proportion in the news title, while the other three elements did not reach 80%, reflecting the fragmentation of short video news. Among the high-frequency words, the frequency of the virtual word “of” was the highest 641 times, and the proportion of monosyllabic words and two-syllable words was as high as 93.2%, which was higher than that of other multi-syllable words, corresponding to the short, concise and interactive characteristics of short video news.

As an emerging form of information dissemination, short video is characterized by short time, concise content and strong interactivity, which brings unprecedented challenges and opportunities to news dissemination. Strengthening interaction, improving content quality, broadening communication channels, and innovating communication forms can improve the quality of short video news dissemination, meet the needs of the audience, and promote the healthy development of the industry.

References

- [1] Tian, Y. (2023). Analysis of new media development opportunities in the context of media integration. *Media and Communication Research*, 4(1), 15-18.
- [2] Zhu, S. (2020). Challenges and access to production and distribution of news in the era of big data. *Proceedings of Business and Economic Studies*, 3(3).
- [3] Sun, Z., & Hou, W. (2022). An analysis of the characteristics and trends of China's mainstream media integration. *Studies in Social Science & Humanities*, 1(5), 58-63.
- [4] Yuan, Y., & Wang, Q. (2024). Characteristics, hotspots, and prospects of short video research: A review of papers published in China from 2012 to 2022. *Heliyon*, 10(3).
- [5] Liu, X., Wu, J., & Zhao, B. (2022, January). Research on the impact of content fragmentation propagation in short videos. In *2021 International Conference on Social Development and Media Communication (SDMC 2021)* (pp. 231-236). Atlantis Press.
- [6] Sun, A. (2022, January). Dynamic modeling of TV news convergence and dissemination under the visual threshold of the internet. In *2022 4th International Conference on Smart Systems and Inventive Technology (ICSSIT)* (pp. 1401-1404). IEEE.
- [7] Li, H. (2023). The language of new media journalism on short video sharing websites: A case study of TikTok journalism. *Academic Journal of Management and Social Sciences*, 3(1), 33-41.
- [8] Li, T. (2024). Media convergence strategies for the interaction between social media and television programs. *Media and Communication Research*, 5(1), 152-158.
- [9] Tang, E. (2020). Analysis of mass communication of short video clips in the era of convergence media. *Sino-US English Teaching*, 17(3), 93-98.
- [10] Du, R. (2023). Expansion of social news communication path under the background of new media integration. *Probe - Media and Communication Studies*, 5(4).
- [11] Sihan, W. (2024). Analysis on the transformation and development path of traditional media based on media convergence. *International Journal of Frontiers in Sociology*, 6(4).
- [12] Yudian, L., Chenhao, L., & Yixin, G. (2021, December). The innovative development of news dissemination in a media convergence environment: Taking Kanghui VLOG as an example. In *2021 3rd International Conference on Economic Management and Cultural Industry (ICEMCI 2021)* (pp. 3410-3413). Atlantis Press.
- [13] Tan, Y. (2023). Analysis of the effectiveness of short video news dissemination based on data mining technology. *Applied Mathematics and Nonlinear Sciences*.
- [14] Zhao, M. (2023). Analysis of the change path of news dissemination under the trend of short-video convergence: Based on multi-priority dynamic thresholding algorithm. *Applied Mathematics and Nonlinear Sciences*, 9(1).
- [15] Wang, Y., & Li, P. (2022). Development and strategy analysis of short video news dissemination under the background of artificial intelligence. *Mobile Information Systems*, 2022(1), 2750925.