IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 3, 2020, accepted November 9, 2020, date of publication November 24, 2020, date of current version January 5, 2021.

Digital Object Identifier 10.1109/ACCESS.2020.3040151

Method for Ranking the Helpfulness of Online Reviews Based on SO-ILES TODIM

HUIFANG DONG, YANHUI HOU, MIN HAO, JIAKUN WANG, AND SHUOSHUO LI
College of Economics and Management, Shandong University of Science and Technology, Qingdao 266590, China

Corresponding author: Yanhui Hou (houyhui@sdust.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51574157, in part by the Natural Science Foundation of Shandong Province under Grant ZR2018MG001, in part by the Qingdao Social Science Planning Project under Grant QDSKL1601107, and in part by the Graduate Science and Technology Innovation Project of Shandong University of Science and Technology under Grant SDKDYC190230.

ABSTRACT Online shopping has become a habit for consumers, who often make purchase decisions based on online reviews. However, the gradual accumulation of reviews has caused an issue associated with information redundancy. Therefore, recommending helpful reviews for consumers has become an urgent problem. Current research on the helpfulness of reviews is mostly at the level of analyzing influencing factors, and few studies have focused on the problem of ranking the helpfulness of reviews. Taking film reviews as the research object, we proposed a SO-ILES TODIM method (a TODIM method based on the intuitive language evaluation set of emotional and ontological features). This method takes into account both semantic indicators (emotional factors and ontological features) and statistical indicators (review length), considers comprehensive information in the review text and has better domain adaptability. First, an intuitive language evaluation set that considers emotional and ontological features was constructed based on statistical rules. Second, a quantitative calculation method that includes an index weight value based on the logit regression model was designed, and it can effectively avoid the subjectivity of the manual assignment method. Finally, based on the degree of membership deviation, the score function and the exact function were designed to realize a ranking of the helpfulness of reviews. Through a case simulation, we show that this method can prioritize reviews that directly evaluate the product. Through a comparative analysis and parameter sensitivity analysis, the stability and scientificity of the SO-ILES TODIM method was demonstrated. This paper broadens the research scope of reviews, enriches the research method of review helpfulness ranking and provides insights for merchants or third-party platforms to manage online reviews.

INDEX TERMS Ranking, helpfulness, online reviews, TODIM.

I. INTRODUCTION

According to data from the National Bureau of Statistics of China, in 2019, China's online retail sales reached 1.06324 trillion yuan, an increase of 16.5% over the previous year. The online penetration rate of online retail sales reached 20.7%, an increase of 2.3 percentage points over the previous year. The continued boom in online shopping has allowed e-commerce systems to accumulate a large number of online reviews, which are an important basis for consumer decision-making. High-quality reviews are effective in helping consumers make purchasing decisions, whereas low-quality reviews waste consumers' time. If reviews are ranked according to their helpfulness and the most helpful information for purchasing decisions is prioritized, then the time cost for consumers to read reviews can be reduced and the efficiency of purchasing decisions can be improved. To investigate this practical problem, this paper studies the ranking problem of the helpfulness of reviews.

At present, research on the helpfulness of reviews mainly focuses on analyzing the influencing factors and constructing prediction models. The objects of study are mainly searched products with few experiential products. Studies have shown that product types can affect consumers' purchasing decisions [1], and research conclusions on search products are not universal for experiential products. Given the current research status, this paper takes movies as an example to rank the helpfulness of experiential product reviews.

The associate editor coordinating the review of this manuscript and approving it for publication was Donato Impedovo.

The review helpfulness ranking problem refers to the ranking of online reviews based on the helpfulness index score. Identifying a method of filtering the helpfulness index of reviews and ranking the helpfulness of reviews are the two areas of emphasis of the research. The following is a review of the research status for these two aspects.

Online reviews mainly include the reviewer's emotional attitude and description of product features. Emotional attitude indicates a like or dislike of the product. Ontological features indicates the consumers' functional preferences for products. Ontology refers to the evaluation objects of reviews. Emotional attitude and ontological features comprehensively cover the semantic information that reviews can convey to readers. Bi *et al.* [2] measured the effects of customer sentiments on customer satisfaction; Bi *et al.* [3], Kauffmann *et al.* [4], Liu *et al.* [5], and Liu *et al.* [6] ranked alternative products using emotion analysis technology; Kumar and Abirami [7] ranked alternative products based on ontological features; Huang and Jiang [8] and Saumya *et al.* [9] calculated the helpfulness score of reviews based on ontological features; and Wang *et al.* [10] realized product ranking by identifying product features and emotional polarity. In addition to semantic information, scholars have constructed statistical indicators to measure the helpfulness of reviews. Singh *et al.* [11] and Shaalan *et al.* [12] built a review ranking model by using information entropy and score distribution. The above studies effectively constructed evaluation indexes of the helpfulness of reviews from semantic and statistical aspects.

Multiattributes of things can reflect the nature of the object. Through multi-criteria decision-making (MCDM) about multiattributes, we can solve many problems, such as the prediction of tourist volume [13], the scheduling of shared bikes [14], management of hotels [15], the evaluation of internet of things platforms [16], and the realization of Importance Performance Analysis (IPA) [17]. The ranking process of reviews is also a MCDM problem, which should take into account the uncertainty of review information, the contradiction between attributes, and the decision maker's loss aversion psychology. Specifically, the uncertainty of review information can be reflected by the size of the attribute value, the contradictory relationship between indicators can be reflected by the size of the attribute weight, and the decision maker's loss aversion can be reflected by the loss attenuation coefficient.

The research emphases of MCDM mainly include two aspects: the construction of evaluation index and the calculation of attribute weight. Wu *et al.* [18] proposed the concept of the hesitant pythagorean fuzzy sets (HPFS) to enhance fuzzy related problems flexibility; Lin *et al.* [19] designed an entropy measurement method to measure the uncertainty of probabilistic language term sets; Wu *et al.* [20] extended VIKOR methods based on the interval type-2 fuzzy best-worst; Liu and Teng [21], Zhang *et al.* [22], Wu and Zhang [23], and Lin *et al.* [24] optimized the MCDM algorithm by constructing attributes and attribute values, which are specifically reflected by the use of the extended probability language TODIM (PL-TODIM) method, fuzzy emotion word framework, intuitionistic fuzzy emotion word framework and probabilistic uncertain linguistic term set; Davoudabadi *et al.* [25], Wang *et al.* [26], and Wu *et al.* [27] optimized the MCDM algorithm via a quantitative weight calculation, which was specifically reflected in the sorting study of alternative schemes by aggregating objective and subjective weights, developing entropy weighting technology and combining network analysis and entropy weight methods, respectively; and Xiao *et al.* [28] developed novel operational laws for a hesitant fuzzy linguistic term set (HFLTS) and applied them to derive the attribute weights. The above research realizes the continuous optimization of the MCDM algorithm.

Previous research has mainly focused on the ranking study of alternative products, which provides a theoretical basis for the ranking study of the reviews on the helpfulness. However, these studies present two deficiencies. First, the evaluation index only considers the emotional factors without considering the ontological characteristics, and since different research objects contain different features, it is necessary to consider ontology features to construct indexes; and second, the weights are calculated by the subjective expert assignment method, which necessitates the design of a quantitative calculation method. In view of the current research status, we improved TODIM method, proposed SO-ILES TODIM, and made up for the two shortcomings of the above research. Taking movies as the research object, we realized the method for ranking the helpfulness of review.

Our contribution includes two aspects: theoretical value and practical significance. The theoretical contribution of this paper is that we propose a SO-ILES TODIM method (a TODIM method based on the intuitive language evaluation set of emotional and ontological features) that takes into account emotional factors and ontology characteristics, makes the evaluation set more applicable in the field, and can use the regression coefficient method to quantify the index weight, thereby avoiding the subjectivity of the manual assignment method. The practical significance of this paper is that the method we proposed can prioritize reviews that directly evaluate the products, thereby reducing the time cost of consumers reading reviews and improving the efficiency of consumers making purchasing decisions based on reviews.

The rest of the paper is organized as follows. Section II introduces the basic concept of TODIM and the research method of this paper. Section III introduces the process of constructing the intuitive language evaluation set of emotional and ontological features (SO-ILES) and the calculation method of attribute weight. Section IV presents the case simulation analysis, comparative analysis, and parameter sensitivity analysis. Section V summarizes the research of this paper.

II. METHODOLOGY

A. BASIC CONCEPTS

Definition 1 [29]: The language evaluation set $S = \{s_\theta \mid \theta = 0, 1, 2, \cdots, 2\,l\}$, $l \in Z^+$, s_θ is an evaluation term indicating the grade of an evaluation index. In the specific field of argument X, if we have $s_{\theta(x)} \epsilon S$, then the intuition language set on X is $T = \left\{\left(x, \left\langle s_{\theta(x)}, \mu(x), \vartheta(x)\right\rangle\right) \mid x \in X\right\}$, where $\mu(x) :\rightarrow [0, 1]$ and $\vartheta(x) :\rightarrow [0, 1]$, which is also known as the attribute value of the evaluation index and presents the membership degree and nonmembership degree of $s_{\theta(x)}$. $t = \left\langle s_{\theta(x)}, \mu(x), \vartheta(x)\right\rangle$ is called the intuitive linguistic number. When $\mu(x) = 1$, the intuitive language set becomes the language evaluation set. For example, $l = 3$, $S = \{s_0 =$ "very bad", $s_1 =$ "quite bad", $s_2 =$ "bad", $s_3 =$ "ordinary", $s_4 =$ "good", $s_5 =$ "quite good", $s_6 =$ "very good"$\}$. In this case, the intuitive language number for $t = \langle s_2, 0.6, 0.3\rangle$ indicates that the probability of an evaluation object belonging to s_2, which is "bad", is 0.6; the probability of not belonging to s_2 is 0.3; and the uncertain probability of an evaluation object is 0.1.

Definition 2: For any two intuitive language numbers $t_1 = \left\langle s_{\theta(t_1)}, \mu\left(t_1\right), \vartheta\left(t_1\right)\right\rangle$ and $t_2 = \left\langle s_{\theta(t_2)}, \mu\left(t_2\right), \vartheta\left(t_2\right)\right\rangle$, the definition score function $F(t)$ and the precise function $G(t)$ are as follows:

$$F(t) = s_{\theta(t)}\left(\frac{\mu(t) - \vartheta(t)}{2}\right)\left(\frac{\mu(t) - \overline{\mu(t)}}{\overline{\mu(t)}}\right) \tag{1}$$

$$G(t) = s_{\theta(t)}\left(\frac{\mu(t) + \vartheta(t)}{2}\right)\left(\frac{\mu(t) - \overline{\mu(t)}}{\overline{\mu(t)}}\right) \tag{2}$$

where $\overline{\mu(t)} = 1/n\sum_1^n \mu(t)$, $n = 1, 2, 3, \ldots$, represents the mean membership degree under the evaluation index. As the frequency of different evaluation indexes in product reviews may differ greatly, a large difference in the membership degree may exist. The score function and accurate function are calculated by using the deviation of membership relative to the mean of membership. Based on the existing research [30], we consulted relevant experts and finally designed formulas (1) and(2).

Definition 3: Any two intuitive language numbers $t_1 = \left\langle s_{\theta(t_1)}, \mu\left(t_1\right), \vartheta\left(t_1\right)\right\rangle$ and $t_2 = \left\langle s_{\theta(t_2)}, \mu\left(t_2\right), \vartheta\left(t_2\right)\right\rangle$ have the following properties:

(1) If $F(t_1) > F(t_2)$, then $t_1 > t_2$;

(2) If $F(t_1) = F(t_2)$,$G(t_1) = G(t_2)$, then $t_1 = t_2$;

(3) If $F(t_1) = F(t_2)$,$G(t_1) > G(t_2)$, then $t_1 > t_2$;

By comparing the score function $F(t)$ and precise function $G(t)$, profit and loss can be qualitatively measured.

Definition 4 [31]: For any two intuitive language numbers $t_1 = \left\langle s_{\theta(t_1)}, \mu\left(t_1\right), \vartheta\left(t_1\right)\right\rangle$ and $t_2 = \left\langle s_{\theta(t_2)}, \mu\left(t_2\right), \vartheta\left(t_2\right)\right\rangle$, the Hamming distance between t_1 and t_2 is as follows:

$$\begin{aligned} d\left(t_1, t_2\right) = &\left|\theta\left(t_1\right)\mu\left(t_1\right) - \theta\left(t_2\right)\mu\left(t_2\right)\right| \\ &+ \left|\theta\left(t_1\right)\left(1 - \vartheta\left(t_1\right)\right) - \theta\left(t_2\right)\left(1 - \vartheta\left(t_2\right)\right)\right| \end{aligned} \tag{3}$$

where $\theta\,(t)$ is function that takes the subscript of avariable.

B. PROBLEM DESCRIPTION

This paper takes film reviews as the research object, selects relevant characteristics as the evaluation index of helpfulness, and determines the ranking of the helpfulness of reviews.

For the sake of uniformity, describe the problem as follows: $N = \{1, 2, 3, \ldots, n\}$, $M = \{1, 2, 3, \ldots, m\}$. $R_i (i \in M)$ is the set of film reviews. The evaluation index set is $C_j (j \in N)$. w_j represent the weight of the jth index, $w_j \geq 0$, and $\sum_{j=1}^n w_j = 1$. $\widetilde{w}_j$ represent the relative weight of the jth index, $\widetilde{w}_j \geq 0$, and $\sum_{j=1}^n \widetilde{w}_j = 1$. To avoid the same ranking problem, set the minimum index of S to 1. Assuming that $l = 2$, the language evaluation set is $S = \{s_1 =$ "very bad", $s_2 =$ "bad", $s_3 =$ "ordinary", $s_4 =$ "good", $s_5 =$ "very good"$\}$. The intuitive language decision matrix is $T = \left[t_{ij}\right]_{m\times n}$. The intuitive linguistic number is $t_{ij} = \left\langle s_{\theta(t_{ij})}, \mu\left(t_{ij}\right), \vartheta\left(t_{ij}\right)\right\rangle$, which represents the evaluation value of the indicator given by the decision-maker to the review R_i under evaluation index C_j, and we also know that $0 \leq \mu\left(t_{ij}\right) \leq 1$, $0 \leq \vartheta\left(t_{ij}\right) \leq 1$ and $0 \leq \mu\left(t_{ij}\right) + \vartheta\left(t_{ij}\right) \leq 1$. $\widetilde{T}$ is a normalized form of T. $\widetilde{t}_{ij}$ is a normalized form of t_{ij}. Finally, the helpful ranking of the review set R is realized according to the intuitive language decision matrix $\widetilde{T} = [\widetilde{t}_{ij}]_{m\times n}$ and the weight vector $\widetilde{w}_j$.

C. RESEARCH METHODS

TODIM is a multiattribute decision making method, which calculates the scheme score through the evaluation set and index weight, and finally achieves the purpose of selecting the best alternative scheme. The main calculation process is as follows: (1) Construct evaluation set. (2) Measure the index weight. (3) Design functions to calculate schemes' score. (4) Select the best alternative scheme according to the score.

In order to rank the helpfulness of reviews, we made three improvements to the TODIM method. The specific manifestations are as follows: (1) Based on emotion analysis and ontological feature model, we propose a new intuitive language evaluation set (SO-ILES). (2) Based on the regression coefficient, we proposed the regression coefficient method, which realized the scientific calculation of the weight value. (3) In order to solve the problem of repeated attribute occurrence in a review, we designed new scoring function and precise function, so that TODIM method can better solve the ranking problem of review helpfulness.

Combined with the calculation process of TODIM, referring to the research of Liu(2019) [21], We designed the calculation process of SO-ILES TODIM as follows:

Step 1: Emotion analysis is carried out for the review text and ontology feature model is constructed for the evaluation object. On this basis, we select evaluation indexes for the helpfulness of reviews, calculate their attribute values and construct an intuitive language evaluation set (SO-ILES) based on emotional and ontological features;

Step 2: Build a normalized decision matrix $\widetilde{T} = [\widetilde{t}_{ij}]_{m\times n}$ based on SO-ILES;

Table 1. Film ontological features and their attentions degree.

Feature	Definition	Attention degree
Story	Describes a film story or plot	0.438
Theme	Describes the theme of the film or reflect the theme of the film	0.258
Character	Describes the actor's performance ability, the role's characteristics, and so on	0.225
Scene	Describes special effects, visual effects, and so on	0.148
Director	Describes the director's level of expertise	0.063

Table 2. Examples of film ontological features.

Feature	Words of feature
Story	a lifetime, a year, a day, an hour 一辈子,一年,一天,一个时辰 (*in Chinese*)
Theme	historical period 历史时期 (*in Chinese*)
Character	Yu Ji, Ba Wang(Xiang Yu), role, actress 虞姬,霸王,角色,演员 (*in Chinese*)
Scene	lights and shadows of film, film editing 光影,剪辑 (*in Chinese*)
Director	Chen Kaige 陈凯歌 (*in Chinese*)

* To save space, only a few feature words are listed for each feature.

Step 3: Build Logistc regression model, using the regression coefficient method to calculate the relative weights of the indexes $\widetilde{w}_j = (\widetilde{w}_1, \widetilde{w}_2, \widetilde{w}_3, \ldots, \widetilde{w}_n)$;

Step 4: Design score function and exact function. Under each evaluation index C_j, compare the score function and exact function of reviews R_i and R_k, obtain the profit-loss analysis matrix, and clearify the advantages and disadvantages of each review's helpfulness under different indexes;

Step 5: Calculate the profit-loss value for each review $\phi_i = (R_i, R_k)$ and construct the profit-loss priority matrix Φ;

$$\Phi = [\phi(R_i, R_k)]_{m\times n} = \sum_i^n \phi_i(R_i, R_k) \tag{4}$$

The calculation formula of $\sum_i^n \phi_i(R_i, R_k)$ is as follows:

$$\phi_i(R_i, R_k) = \begin{cases} \sqrt{\dfrac{\widetilde{w}_j d(\widetilde{t}_{ij}, \widetilde{t}_{kj})}{\sum_{j=1}^n \widetilde{w}_j}}, & s(\widetilde{t}_{ij}) > s(\widetilde{t}_{kj}) \\ 0, & s(\widetilde{t}_{ij}) = s(\widetilde{t}_{kj}) \\ -\dfrac{1}{\delta}\sqrt{\dfrac{\sum_{j=1}^n \widetilde{w}_j d(\widetilde{t}_{ij}, \widetilde{t}_{kj})}{\widetilde{w}_j}}, & s(\widetilde{t}_{ij}) < s(\widetilde{t}_{kj}) \end{cases} \tag{5}$$

where $i, k \in M, j \in N$. δ is the loss attenuation coefficient, $0 < \delta < \frac{\sum_{j=1}^n \widetilde{w}_j}{\widetilde{w}_j}$.

Step 6: Under the evaluation index of $C_j (j \in N)$, aggregate the profit-loss priority matrix to compute the overall priority $Z(R_i)$; then, standardize the overall priority $Z(R_i)$, calculate the normalized priority $\Gamma(R_i)$; and rank the reviews by normalized priority, with a greater $\Gamma(R_i)$ value indicating higher ranking of R_i.

Aggregation formula is as follows:

$$Z(R_i) = \sum_{j=1}^n \Phi(R_i, R_k) \quad i, k \in M \tag{6}$$

Normalization formula is as follows:

$$\Gamma(R_i) = \frac{Z(R_i) - \min_i \{Z(R_i)\}}{\max_i \{Z(R_i)\} - \min_i \{Z(R_i)\}} \tag{7}$$

where $i, k \in M$ and $0 \leq \Gamma(R_i) \leq 1$.

III. SO-ILES TODIM

For the ranking of the helpfulness of reviews, three points need to be considered: the selection of the evaluation index of the helpfulness of reviews, the calculation of the index attribute value, and the calculation of the index weight. We design an intuitive language evaluation set based on emotional and ontology features (SO-ILES) to realize the selection of evaluation indexes and the calculation of attribute values. This method has better domain adaptability and contains comprehensive review text information. In addition, we design a quantitative calculation method of the index weight value based on the logit regression model, which avoids the subjectivity of manual assignment.

A. SO-ILES

1) SELECTION OF EVALUATION INDEX

The selection of evaluation index for the helpfulness of reviews is the basis for the realization of review helpfulness ranking. The selection process has the following steps. First, the literature on the influencing factors of the helpfulness of published reviews [1], [11], [32]–[37] was evaluated. Second, we discussed screen evaluation indexes with relevant experts based on existing research results. Through the literature review, it was found that the emotional attitude, frequency of emotional words, and frequency of feature words contained in reviews will affect the helpfulness of reviews, among which emotional words include positive emotional words, negative emotional words, and neutral emotional words. Because neutral emotional words have a neutral emotional attitude and are not persuasive to consumers, this paper will not include these words in the study. With regard to feature words and depending on the research object, different research objects have different feature words. Based on the author's previous research [38], film features were selected according to the concept model of film ontology. Table 1 lists the research results of film ontological features and consumers' attention to film features in the research [38]. Table 2 lists some examples of film ontological features. Due to the differences in language and culture, wherever film reviews are involved in

Table 3. Evaluation indexes and their definitions.

Evaluation indicator			
Primary indicator	Secondary indicator	Mark	Definition
Emotional consistency		Senti_cons	Consistency in emotional attitudes in reviews
Emotional words (PN)	Positive words	P_n	Number of positive emotion words included in the review
	Negative words	N_n	Number of negative emotion words included in the review
Film features (TZ)	Story	Story	Number of words that indicate the plot of a film in a review
	Theme	Theme	Number of words that indicate the theme of a film in a review
	Character	Character	Number of words that indicate character (or actor) of a film in a review
	Scene	Scene	Number of words that indicate scene (editing, etc.) of a film in a review
	Director	Director	Number of words that indicate director of a film in a review
Emotional intensity		Intensity	Number of adverbs included in the review
Length of review		Len	Total number of words included in the review

this paper, we provide two forms: one is the corresponding English translation and the other is the Chinese translation.

Emotional intensity reflects the intensity of emotional attitudes and may also affect the helpfulness of reviews. Although consistent conclusions have not been obtained as to whether emotional intensity has a significant effect on the helpfulness of reviews, indicators of emotional intensity were added to improve our study. In addition, a large number of studies [35]–[37] have demonstrated that the length of reviews affects their helpfulness. Due to the lack of statistical indicators in this paper, the review length index was added for research. Table 3 summarizes the review helpfulness evaluation indexes and their definitions.

2) CALCULATION OF ATTRIBUTE VALUES

The attribute value of the index includes the membership degree and nonmembership degree, which reflects the relationship between the index and the language evaluation value.

a: EMOTIONAL CONSISTENCY (SENTI_CONS)

Emotional consistency refers to the degree of consistency in the emotional attitude of the review. Previous studies usually only use linear addition and subtraction of emotion words to obtain the emotional tendency of the review, which is divided into three categories: positive, negative, and neutral. This approach ignores the semantic emphasis of language art. In language art, the key point that people want to express is behind the transition words and methods that use linear addition and subtraction will often offset the importance of the emotional attitude. This paper introduces the emotional consistency of the attention factor to calculate review helpfulness. Because neutral emotion words do not influence the consistency of emotion, this paper does not calculate neutral words; instead, only positive and negative emotion words are studied.

The emotional consistency of the membership formula of the ith review is as follows:

$$Senti\_cons\_\mu_i = \left| \frac{\alpha P\_n_i - (1-\alpha) N\_n_i}{\alpha P\_n_i + (1-\alpha) N\_n_i + \varepsilon} \right| \tag{8}$$

In addition to emotion words, reviews do contain other words. We use other words in the proportion of the total number of words to measure the emotional consistency of nonmembership degrees. The emotional consistency of the nonmembership degree formula of the ith review is as follows:

$$Senti\_cons\_\vartheta_i = (1 - Senti\_cons\_\mu_i) \times \left(1 - \frac{P\_n_i + N\_n_i}{Q_i}\right) \tag{9}$$

where $i \in M$, $Senti\_cons\_\mu_i$ is the membership degree of emotional consistency in the ith review; $Senti\_cons\_\vartheta_i$ is the nonmembership degree of emotional consistency in the ith review; P_n_i is the number of positive words in the ith review; N_n_i is the number of negative words in the ith review; Q_n_i is the number of total words in the ith review. We set ε to be an infinitesimal constant to prevent the divisor from being zero. Let $\varepsilon = 0.001$. α represents the coefficient of attention in the ith review, with $0 \le \alpha \le 1$. A higher value of α, corresponds to a review with a greater emphasis on emotion. When the emotion of the review was consistent, the attention to positive and negative words was the same; therefore, the attention coefficient was set to 0.5. When the review emotion was inconsistent, the attention to positive and negative words was different, and the attention coefficient ranged from $0.5 < \alpha \le 1$. Alpha values within this range represent weighted attention to emotion, except that the higher the alpha value, the higher the attention to emotion. For example, if $\alpha = 1$, then the review emphasizes only one emotion, either positive or negative. In this paper, the attention coefficient is set to 0.7 after discussion with experts. In addition, because accent semantics in reviews cannot be automatically identified, manual annotation is required during text processing.

b: EMOTIONAL WORDS (PN)

Use the proportion of positive words in all emotional words to indicate the membership degree of positive words. The membership degree formula of positive words in the ith review is as follows:

$$P\_n\_\mu_i = \frac{P\_n_i}{P\_n_i + N\_n_i + \varepsilon} \tag{10}$$

The nonmembership degree is indicated by the proportion of other words in the total number of words in the reviews. The nonmembership degree formula of positive words in the ith review is as follows:

$$P\_n\_\vartheta_i = (1 - P\_n\_\mu_i)\left(1 - \frac{P\_n_i}{Q_i}\right) \tag{11}$$

where $i \in M$, $P\_n\_\mu_i$ is the membership degree of positive words in the ith review; $P\_n\_\vartheta_i$ is the nonmembership degree of positive words in the ith review.

The attribute values of negative words are calculated in the same way.Let the membership degree of the negative words in ith review be $N\_n\_\mu_i$ and the nonmembership degree of the negative words in ith review be $N\_n\_\vartheta_i$.

$$N\_n\_\mu_i = \frac{N\_n_i}{P\_n_i + N\_n_i + \varepsilon} \tag{12}$$

$$N\_n\_\vartheta_i = (1 - N\_n\_\mu_i)\left(1 - \frac{N\_n_i}{Q_i}\right) \tag{13}$$

where $i \in M$.

c: FILM FEATURES (TZ)

This paper contains five types of film features, and the calculation process of index attribute values under different features is the same. The importance of film features to the helpfulness of the review is adjusted by weight in the following text. Due to the peculiarities of film reviews, the same feature may be mentioned many times in the same review. For example, a review contains multiple actors' names and mentions multiple storylines. Therefore, the membership degree of film features can be calculated by frequency. Only when film features are matched with emotional words can the critic's attitude be expressed. Therefore, the membership degree is calculated by the proportion of feature words in emotional words. The membership degree of the film features of nth category in ith review is as follows:

$$TZ\_\mu_{in} = \frac{TZ_{in}}{P\_n_i + N\_n_i + TZ_{in} + \varepsilon} \tag{14}$$

When the film features are not modified by emotional words, these features only express the objective statement of the critic but do not convey an emotional attitude to the reader. However, the features can convey some objective information about the film. Therefore, the proportion of film features in the remaining words is used to express the degree of nonmembership. The nonmembership degree of the film features of nth category in ith review is as follows:

$$TZ\_\vartheta_{in} = (1 - TZ\_\mu_{in}) \times \frac{TZ_{in}}{P\_n_i + N\_n_i + + \sum_1^n TZ_i + \varepsilon} \tag{15}$$

where $i \in M$, TZ_μ_{in} is the membership degree of the film features of nth category in ith review; TZ_ϑ_{in} is the nonmembership degree of the film features of nth category in ith review; TZ_{in} is number of film features of the nth category included in the ith review; $n = 1, 2, 3, 4, 5$ represents the film ontological features in Table 1; that is, the story, theme, character, scene, and director, respectively; and TZ_i is the total number of feature words included in the ith review.

d: EMOTIONAL INTENSITY (INTENSITY)

Emotional intensity reflects the intensity of the critic's emotional tendency. For example, for "Z's acting is good" and "Z's acting is very good", the reference value for readers is different. The modifier effect of emotional intensity words enhances the credibility of the review information. In addition, the attribute value of emotional intensity words will not change because of different modifiers; that is, the intensity modifies semantic objects to the same degree. Therefore, according to the method of Zhang (2020) [39], this paper divides emotional intensity into five levels and sets the attribute value in the same way. The specific meanings are shown in Table 4.

Table 4. Examples of emotional intensity levels and their attribute values.

Emotional intensity levels	Words of emotional intensity	Attribute values
1	extremely (*in Chinese* 极其)	<0.9,0.05>
2	very (*in Chinese* 很)	<0.7,0.25>
3	more (*in Chinese* 较)	<0.5,0.4>
4	a bit (*in Chinese* 稍)	<0.25,0.7>
5	a little less (*in Chinese* 欠)	<0.05,0.9>

* To save space, only one emotional intensity word is listed for each level.

Considering that a review may contain multiple emotional intensity words at the same time, the average emotional intensity value is used to represent the emotional intensity value of the whole sentence. We set the membership degree of emotional intensity as Adv_u_i and the nonmembership degree of emotional intensity as Adv_v_i. The calculation methods are as follows:

$$\mathrm{Adv}\_\ u_i = \frac{\mathrm{Adv}_p}{m_i} \tag{16}$$

$$Adv\_v_i = 1 - \frac{Adv_p}{m_i} \tag{17}$$

where $i \in M$, Adv_p represents the sum of attribute values corresponding to different levels of emotional intensity words; $p = 1, 2, 3, 4, 5$ represents the emotional intensity levels in Table 4; and m_i is the number of emotional intensity words included in the ith review.

e: LENGTH OF REVIEWS (Len)

The review length is defined as the total number of words contained in the review. Theoretically, the more words will contain more information and the review will be more helpful. However, if the number of review words is too high, then the information will be redundant and the helpfulness of the review will decrease. In addition, reviews that contain more incongruent emotions tend to be longer. Therefore, we calculate the deviation degree and the membership degree of emotional consistency from the average review length. The membership degree of review length in ith review is as follows:

$$\mathrm{Len\_}\,\mu_i = \mathrm{Senti\_\,cons\_}\,\mu_i \times \sqrt{\left|\frac{\mathrm{Len}_i - \frac{1}{M}\sum_{i=1}^{M}\mathrm{Len}_i}{\frac{1}{M}\sum_{i=1}^{M}\mathrm{Len}_i}\right|} \tag{18}$$

The nonmembership degree of review length in ith review is as follows:

$$\mathrm{Len\_}\,\vartheta_i = \mathrm{Senti\_\,cons\_}\,\vartheta_i\,(1 - \mathrm{Len\_}\,\mu_i) \tag{19}$$

where $i \in M$, M is the total number of reviews; $\mathrm{Len\_}\,\mu_i$ is the membership degree of review length in ith review; $\mathrm{Len\_}\,\vartheta_i$ is the nonmembership degree of review length in ith review; Len_i is the total number of words contained in the review.

3) INTUITIVE LANGUAGE EVALUATION SET

The attribute values of each index are calculated according to formulas 8-19. Let $l = 2$; then, the language assessment set S is $\{s_1 =$ "very bad", $s_2 =$ "bad", $s_3 =$ "ordinary", $s_4 =$ "good", $s_5 =$ "very good"$\}$. According to Liu (2019) [21], we assign a value to S. S is defined as $\{0 \leq s_1 \leq 0.2, 0.2 < s_2 \leq 0.4, 0.4 < s_3 \leq 0.6, 0.6 < s_4 \leq 0.8, 0.8 < s_5 \leq 1\}$. By combining the language evaluation set with each attribute value, the intuitive language evaluational set (SO-ILES) based on emotional and ontological features, also known as the decision matrix T is obtained.

After obtaining the decision matrix $T = \left[t_{ij}\right]_{m\times n}$, it is necessary to normalize the matrix, which can avoid the impact of different data dimensions on the decision results. Decision indicators are usually divided into cost (denoted as Cost) and benefit (denoted as Benefit). Referring to the method of Lin (2019) [40], standardized processing is carried out according to formula (20) to obtain the standardized decision matrix $\tilde{T} = [\tilde{t}_{ij}]_{m\times n}$.

$$\widetilde{t_{ij}} = \begin{cases} s_\theta\left(t_{ij}\right), & j \in Cost \\ s_{2l-\theta}\left(t_{ij}\right), & j \in Benefit \end{cases} \tag{20}$$

where $i \in M$, $\widetilde{t_{ij}}$ is the intuitive linguistic number after standardizing; $s_\theta \in S$, $s_{2\,l-\theta} \in S$; $\tilde{T}$ is the standardized decision matrix.

B. CALCULATION OF EVALUATION INDEX WEIGHT

To avoid the subjectivity of manual weight assignments, index weights are standardized based on regression model coefficients. The coefficient of the regression model reflects the influence degree of the independent variable on the dependent variable and can be used to measure the weight of the index. Because the helpful classification of reviews is a binary classification, we builds a logit regression model based on 300 reviews about popular movies from the Douban website obtained by web crawler technology. The regression coefficient is shown in Table 5. We calculate the index weight on this basis.

According to the section of "SELECTION OF EVALUATION INDEX", the impact of factors at the emotional and feature level on the helpfulness of reviews is obvious. To ensure the scientific nature of the model, only indicators at the emotional and feature level are selected to construct the model. The weight of other indicators is adjusted through the influence coefficient λ $(0 \leq \lambda \leq 1)$. Given the greater impact of indicators at the emotional and feature levels on the helpfulness of reviews, the total influence coefficient of the two levels λ was set as 0.9 after discussion with experts. At present, unanimous conclusions have not been reached on whether emotional intensity affects the helpfulness of reviews; therefore, it is not considered. In addition, we assumed that the distribution of index weights is consistent when the data have the same source. The data in this paper are all from the Douban movie website; therefore, the index weight value obtained in this section applies to the case analysis in the following article.

Table 5. Regression coefficients in the Logit model.

Helpfulness	Coef.	P>|z|
Senti_cons	0.961	0.002
P_n	0.285	0.001
N_n	0.233	0.010
TZ	0.112	0.001

The weights of each indicator in Table 5 $\dot{w}_a$ is defined as follows:

$$\dot{w}_a = \lambda \frac{w_a}{\sum_{a=1}^{4} w_a} \tag{21}$$

where $\dot{w}_a$ is the weight of each indicator in the study; $w_a$is the regression coefficient of the factors in the logit model,$a = 1, 2, 3, 4$, and they represent each influencing factor in Table 5, $w_1 = 0.961$, the rest are assigned in order; λ is the influence coefficient, and its value is 0.9.

$a = 4$ is the weight of the film features (TZ), which is calculated as 0.063 by formula (21). It is worth noting that film features include director, character and other features. The degree of influence of different features on moviegoers is different. For example, some people like a certain theme, the reviews on the theme items have more information value for such people; therefore, it is necessary to subdivide film features (TZ). The film features (TZ) can be described by the film ontology model [38]. As shown in Table 1, we calculate the weight of film features based on an attention degree. The formula of feature weight $\dot{w}_b$ for the film is shown in Equation (22):

$$\dot{w}_b = \dot{w}_4 \frac{w_b}{\sum_{b=1}^{5} w_b} \tag{22}$$

where $\dot{w}_b$ is the weights of each film feature; $\dot{w}_4$ is the weight of film features (TZ); w_b is the attention degree of each film feature, $b = 1, 2, 3, 4, 5$, which represents the features of a film in Table 1. $w_1 = 0.438$, the rest are assigned in order.

Emotional intensity reflects the intensity of expressed emotions. The length of the text reflects the richness of review information. These indicators were added to improve the study, and the relationship between them was adjusted by the weight coefficient γ $(0 \leq \gamma \leq 1)$. Since most of the evaluation indicators in this paper are at the semantic level, after discussion with relevant experts, the length of the review is given greater weight. We set $\gamma = 0.4$. The method of weight calculation for emotional intensity $\dot{w}_c$ and review length $\dot{w}_d$ is shown in formulas 23-24:

$$\dot{w}_c = \gamma \left(1 - \sum\nolimits_{a=1}^{4} \dot{w}_a\right) \tag{23}$$

$$\dot{w}_d = (1 - \gamma) \left(1 - \sum\nolimits_{a=1}^{4} \dot{w}_a\right) \tag{24}$$

where $\dot{w}_a$ is the weight of each indicator in the study, it is calculated according to formula (21).

Table 6 is a summary of the weights, which are calculated according to the formulas 21-24.

Table 6. Summary table of weights.

Evaluation index	Weight
Senti_cons	0.544
P_n	0.161
N_n	0.131
Story	0.024
Theme	0.014
Character	0.013
Scene	0.008
Director	0.003
Intensity	0.040
Len	0.061

Finally, we normalize the index weight and calculate the relative weight $\widetilde{w}_j$ of C_j by formula (25).

$$\widetilde{w}_j = \frac{\dot{w}_j}{\max_{1 \leq j \leq n} \dot{w}_j} \quad j \in N \tag{25}$$

where $\dot{w}_j$is the weight of each review's helpfulness evaluation index as shown in Table 6.

IV. CASE ANALYSIS

To verify the validity of the SO-ILES TODIM method, a case analysis, comparative analysis, and parameter sensitivity analysis were carried out with film short review data as the research object.

A. DATA ACQUISITION AND PROCESSING

We choose the Douban film website (the largest film review platform in China) as the data source for this paper and select the classic movie *Farewell My Concubine* (Starring Leslie Cheung) for case analysis. As is well known, there are some admittedly rational reviews. By referring to these rational reviews, we can judge whether the result of the experiment is reliable. In addition, a film often contains many reviews, and it may be impossible for a reader to read all of them. The first page of reviews on third-party websites is the easiest and first to be seen by readers. Considering readers' reading habits and time costs, only the first page of reviews on the website is selected for the case analysis, which contains twenty reviews in all. Python was used to preprocess text, such as punctuation removal and word segmentation, and to implement the programming, such as the builder program of SO-ILES and the algorithm program of SO-ILES TODIM.

Data source: $https://movie.douban.com/subject/1$ $291546/comments?sort = new\_score\&status = P$

Because pages on the site are constantly updated, the current order of reviews may be different from the order of reviews in the experiment. To ensure the credibility of the data, a screenshot was taken to save the data, as shown in Figure 1. To facilitate the presentation, the reviews were marked manually and marked as "rank_n", $n = 1, 2, 3, \ldots, 20$, as shown in the red box in Figure 1.

B. SORTING PROCESS

Step 1 (Construct the Intuitive Language Evaluation Set (SO-ILES) Based on Emotional and Ontological Features): Formulas 8-19 were used to calculate the values of the evaluation index of the review helpfulness. The values were combined with the language evaluation set to construct SO-ILES. Because of space limitation, only SO-ILES values under the five indicators of the first five reviews are listed in Table 7.

Step 2 (Build the Normalized Decision Matrix $\tilde{T}$): SO-ILES is called the decision matrix T. Formula (20) is used to obtain the normalized decision matrix $\tilde{T}$. There is no doubt that when the emotional consistency, emotional tendency, and emotional intensity in film reviews are more obvious, then readers are more easily persuaded. A longer review corresponds to a greater number of characteristics, a more informative text, and a more perceptive review. Therefore, the research indicators in this paper are all benefit indicators, that is, the higher the value, the more helpful the reviews is. It is worth noting that since the research object is short film reviews, the length of the review is limited by both third-party websites and consumer review habits; in this case, a longer review corresponds to a greater likelihood of containing more helpful information. By formula (20), we know that$\tilde{T} = T = [t_{ij}]_{m \times n}$.

Step 3 (Calculate the Relative Weights of Indicators $\tilde{w}$): According to the section "CALCULATION OF EVALUATION INDEX WEIGHT", formulas 21-25 are used to calculate the relative weight of each evaluation index. After calculation, we obtained the relative weight as follows: $\tilde{w} = (1, 0.299, 0.241, 0.044, 0.026, 0.024, 0.015, 0.006, 0.07$ $4, 0.112)$. The sum of the relative weights is $\sum \tilde{w} = 1.841$.

Step 4 (Calculate the Score Function $F(t)$ and the Exact Function $G(t)$ and Carry Out the Profit-Loss Analysis): Based on formulas 1-2 and definition 3, we can compare the

Figure 1. **Screenshot of reviews on Douban's website.**

profit-loss of each review. For the convenience of expression, if the comparison result is "greater than", then the relationship is marked as "1", which indicates that the review is superior to the other; if the comparison result is "equal", then the relationship is marked as "0", which indicates that the two reviews are equivalent in terms of the helpfulness of the reviews; and if the comparison result is "less than", then the relationship is marked as "−1", which indicates that the review is less helpful than the other review. According to formula(3), the specific merits and demerits are calculated and their magnitude is d, thus, the degree of merits and demerits of each review can be quantitatively understood.

Because of space limitations, only the profit-loss analysis of the first five reviews in "Senti_cons" is listed as shown in Tables 8 and 9. Taking the cell in the first row and the second column as an example, the score function values

Table 7. SO-ILES (Part).

	Senti_cons	P_n	N_n	Story	Theme
R_1	$\langle s_2, 0.200, 0.677\rangle$	$\langle s_3, 0.500, 0.462\rangle$	$\langle s_3, 0.500, 0.462\rangle$	$\langle s_1, 0.000, 0.000\rangle$	$\langle s_1, 0.000, 0.000\rangle$
R_2	$\langle s_4, 0.000, 1.000\rangle$	$\langle s_1, 0.000, 1.000\rangle$	$\langle s_1, 0.000, 1.000\rangle$	$\langle s_5, 1.000, 0.000\rangle$	$\langle s_1, 0.000, 0.000\rangle$
R_3	$\langle s_5, 0.000, 1.000\rangle$	$\langle s_1, 0.000, 1.000\rangle$	$\langle s_1, 0.000, 1.000\rangle$	$\langle s_1, 0.000, 0.000\rangle$	$\langle s_1, 0.000, 0.000\rangle$
R_4	$\langle s_5, 0.495, 0.750\rangle$	$\langle s_4, 0.750, 0.231\rangle$	$\langle s_2, 0.250, 0.731\rangle$	$\langle s_1, 0.000, 0.000\rangle$	$\langle s_1, 0.000, 0.000\rangle$
R_5	$\langle s_3, 0.300, 0.646\rangle$	$\langle s_1, 0.000, 1.000\rangle$	$\langle s_5, 0.999, 0.001\rangle$	$\langle s_1, 0.000, 0.000\rangle$	$\langle s_1, 0.000, 0.000\rangle$

Table 8. Analysis of profit - loss relationship.

	R_1	R_2	R_3	R_4	R_5
R_1	0	-1	-1	1	-1
R_2	1	0	0	1	1
R_3	1	0	0	1	1
R_4	-1	-1	-1	0	-1
R_5	1	-1	-1	1	0

Table 9. The value of profit - loss analysis.

	R_1	R_2	R_3	R_4	R_5
R_1	0.000	0.523	0.523	2.341	0.784
R_2	0.523	0.000	0.000	2.864	1.307
R_3	0.523	0.000	0.000	2.864	1.307
R_4	2.341	2.864	2.864	0.000	1.558
R_5	0.784	1.307	1.307	1.558	0.000

of R_1 and R_2 are calculated by formula (1) as 0.217 and 0.5, respectively. Since $F(R_1) < F(R_2)$, the relationship between R_1 and R_2 is "-1". The result proves that R_1 is less helpful than R_2, and the degree of its disadvantage is 0.523.

Step 5 (Calculate the Priority of Each Review Under Each Evaluation Index to Build a Profit-Loss Priority Matrix $\Phi_{20\times 20}$): According to formulas 4-5, we set the loss attenuation coefficient $\delta = 1$ and calculate the priority of reviews under each evaluation indicator ϕ_i to build the profit-loss priority matrix Φ. Following up from the previous section, we also listed the priority values of the first five reviews under the "Senti_cons" indicator. As shown in Table 10, take the first row and second column cell as an example, $\phi_1(R_1, R_2) = -(1.841 * 0.523/1)^{1/2}$. Because the report retains only three decimal places, the manual calculation of $\phi_1(R_1, R_2)$ and program calculation of the numerical result will be slightly different in Table 10.

Table 10. Priority values (Part).

	R_1	R_2	R_3	R_4	R_5
R_1	0.000	-0.980	-0.980	1.129	-1.200
R_2	0.534	0.000	0.000	1.249	0.843
R_3	0.534	0.000	0.000	1.249	0.843
R_4	-2.074	-2.293	-2.293	0.000	-1.691
R_5	0.653	-1.549	-1.549	0.921	0.000

Step 6 (Rank the Helpfulness of Reviews According to Their Overall Priority): According to formula (6), we obtain the overall priority Z. According to formula (7), the overall priority Z is normalized and then the final ranking is obtained according to the normalized value.

In step 5, we obtain the priority matrix $\Phi_{20\times 20}$. We take the first review as an example $\Phi(R_1, R_k) = \{0.000, -2.054, -1.550, 1.684, 0.694, 3.803, 5.171, 0.954, 0.655, 2.777, -7.204, 0.652, -7.204, 0.950, 6.881, 0.159, -4.991, -1.093, 4.849, 4.270\}$, where $k = 1, 2, 3, \ldots, 20$. $\Phi(R_1, R_k)$ represents the priority of the first review relative to the other 20 reviews. For example, a value of 0 compared with its priority indicates that the helpfulness of its review is neither better nor worse than itself. By formula (6), the review helpfulness priority is aggregated, and then we obtain the overall priority $Z(R_1) = 3.849$.

The overall priority for all reviews is $Z = \{3.849, -8.670, -34.469, -67.390, -65.634, -156.586, -153.876, -68.576, -34.803, -42.841, 110.902, -80.349, 110.902, -122.327, -260.785, -89.593, 34.480, -61.967, -168.807, -108.159\}$. The normalized priority of all reviews is $\Gamma = \{0.712, 0.678, 0.609, 0.520, 0.525, 0.280, 0.288, 0.517, 0.608, 0.586, 1.000, 0.485, 1.000, 0.373, 0.000, 0.461, 0.794, 0.535, 0.247, 0.411\}$. In descending order, the helpfulness of the 20 reviews is $R_{11} \succcurlyeq R_{13} \succ R_{17} \succ R_1 \succ R_2 \succ R_3 \succ R_9 \succ R_{10} \succ R_{18} \succ R_5 \succ R_4 \succ R_8 \succ R_{12} \succ R_{16} \succ R_{20} \succ R_{14} \succ R_7 \succ R_6 \succ R_{19} \succ R_{15}$. The reviews are subtitled as the original ranking of Douban's website.

To understand the analysis of the results in the following part, this paper defines "direct evaluation of a film" as a review for a particular film and in which the specific name of the movie can be determined and "indirect evaluation of a film" as a review that applies to all films in general and in which the specific name of the movie cannot be determined. There is no doubt that when we are faced with a new product, we tend to obtain specific information. It is easy to understand that reviews are more helpful when they refer specifically to a film. In other words, when reviews refer to films in general, the helpfulness of reviews is relatively low.

The experimental result shows that the ranking results obtained by using the SO-ILES TODIM method are different from those obtained by using the website. The number one review on the site was ranked number four in the results of this paper. The first review on the site read as follows: " Kaige Chen can live on it twice. Now it seems to be excusable that he is at his wit's end. (*in Chinese:* 陈凯歌可以靠它吃两辈子饭了，现在看来江郎才尽也情有可原。)". This review expresses

Table 11. The top 5 reviews in SO-ILES TODIM_rank and Net_rank and their corresponding text.

Rank	SO-ILES TODIM_rank	Net_rank
1	**Dieyi Cheng and the protagonist of** *The Legend of 1900*, **are of the same kind.** 程蝶衣，和《海上钢琴师》的主角，都是一类人。(*in Chinese*)	Kaige Chen can live on it twice. Now it seems to be excusable that he is at his wit's end. 陈凯歌可以靠它吃两辈子饭了，现在看来江郎才尽也情有可原。(*in Chinese*)
2	**...Such a pathetic Dieyi makes people heartbreaking. He is obsessed and trapped too deep into the play. So is the reality of him.** ……如此悲哀可怜的蝶衣让人心痛，不疯魔不成活，人戏不分，现实的他也是如此。(*in Chinese*)	**We once agreed a lifetime. Even it is less a year, a day, an hour, but it still cannot be seen as a whole lifetime.** 说好了一辈子，少一年、一天、一个时辰，都不是一辈子。(*in Chinese*)
3	Such a gorgeous national treasure··· If the story takes place in the Republic of China, this period is usually the most difficult to pass. 那么好的国粹……大多数开始于民国间的故事，最难捱的都是那段时间。(*in Chinese*)	**Telling the whole truth in front of everyone, he said, the Consort Yu is the real Consort Yu, but the Ba Wang Xiang Yu is the fake one.** 他竟当面一语点破：虞姬是真虞姬，霸王是假霸王。(*in Chinese*)
4	**The obsession with the stage carries over into everyday life.** 人戏不分，不疯魔不成活。(*in Chinese*)	With this movie, I'm willing to forgive all of Chen Kaige's stinker movies...With this film, Leslie Cheung is a legend in my heart forever. 就凭这个，我愿意原谅陈凯歌一切的烂片……就凭这个，哥哥你是我心中永远不朽的传奇。(*in Chinese*)
5	**He is really obsessed.** 不疯魔，不成活。(*in Chinese*)	**In front of the non-professional Leslie Cheung, Fengyi Zhang, who was trained by the college, seems so dim.** 在野路子出身的张国荣面前，学院出身的张丰毅显得那么单薄。(*in Chinese*)

* Bold reviews indicate direct evaluation review on the film.

only the professional level of the director, which is an indirect evaluation of a film. With relatively low helpfulness, its review ranking has dropped. The first review in the results of this paper is as follows: " Dieyi Cheng and the protagonist of *The Legend of 1900*, are of the same kind. (*in Chinese:* 程蝶衣和,《海上钢琴师》的主角,都是一类人。)". Although it was ranked 11th on the original site, it is a direct evaluation of a film, its review ranking has increased. The analysis results show that the SO-ILES TODIM method can prioritize to the reviews that directly and specifically describe a film, which is more helpful to purchasing decisions. Therefore, the SO-ILES TODIM method is more scientific.

To avoid chance, we expand the comparison to the top five. The sorting result of the website is denoted as "Net_rank" and the sorting result of this paper is denoted as "SO-ILES TODIM_rank". To facilitate the presentation, the longer review text is manually omitted and replaced by an ellipsis, and it does not affect the judgment of the helpfulness value of the review. Table 11 shows that SO-ILES TODIM_rank contains four direct evaluation reviews and one indirect evaluation review while Net_rank contains three direct evaluation reviews of the film and two indirect evaluation reviews. Since SO-ILES TODIM_rank has more direct evaluation reviews than Net_rank and the ratings of those reviews are relatively near the top, they both are conducive to reducing the time cost of readers and beneficial for consumers to make purchasing decisions. Therefore, the effectiveness and superiority of the SO-ILES TODIM method are verified.

The amount of experimental data is relatively small, thus, the difference with the original website is relatively small. However, e-commerce systems have tens of thousands of reviews. These differences can be magnified by the large volume of data. The review ranking implemented by the SO-ILES TODIM method will prioritize presenting a direct evaluation review of a film and realizing the helpfulness ranking of reviews. It can greatly reduce the redundancy of review information and improve the efficiency of consumers' purchasing decisions.

C. COMPARATIVE ANALYSIS

Controversy remains over whether to select the index of emotional intensity in section III. Thus, although most scholars have studied the influence of emotional intensity on the helpfulness of reviews, the research conclusions are not uniform. To further study whether emotional intensity has an impact on review ranking, this paper excluded the index of emotional intensity to perform a comparative analysis of the helpfulness ranking of reviews again.

According to the processing of SO-ILES TODIM, when the indicator of emotional intensity is excluded, the weight allocation of each indicator will be affected. In this paper, the weight of each index is adjusted by the proportional allocation method. Then, we calculate the value of profit-loss priority and construct the profit-loss priority matrix. The weight adjustment formula is as follows:

$$w\_no\_adv_j = w_j + w_{adv} \frac{w_j}{\sum_{adv \notin C_j} w_j} \tag{26}$$

where $j \in N$, adv is the the emotional intensity index; C_j is the set of evaluation indexes; w_j is the weight of each indicator when considering emotional intensity and does not include the emotional intensity index; w_{adv} is the weight of the emotional intensity index, which is 0.04 as shown in Table 6; $w\_no\_adv_j$ represents the weight of each indicator regardless of the emotional intensity.

Defining $\widetilde{w\_no\_adv}$ as the relative weight after adjustment, which is calculated by formula (25). $\widetilde{w\_no\_adv}$ =

(1.0, 0.296, 0.241, 0.044, 0.026, 0.024, $\underbrace{0.015}$, 0.006, 0.112). The sum of the relative weights is $\sum w\_no\_adv = 1.764$.

For comparison with the information in section "SORTING PROCESS", the first five calculated values of reviews under the "Senti_cons" indicator are also listed here. According to steps 4-6, the priority matrix is calculated as $\acute{\Phi}_{20\times 20}$. Taking the first review as an example, $\acute{\Phi}(R_1, R_k) = \{0.000, -4.460, -4.460, 0.755, -1.410, 1.733, 5.140, 1.894, 0.029, -2.201, -6.719, 3.026, -6.719, 2.201, 5.416, -2.202, -4.461, 1.449, 3.589, 4.808\}$, where $k = 1, 2, 3, \ldots, 20$. $\acute{\Phi}(R_1, R_k)$ represents the priority of the first review relative to the other 20 reviews. By formula (6), we gather to review helpfulness priorities, and then we obtain the overall priority $\acute{Z}(R_1) = -6.994$.

Regardless of emotional intensity, the overall priority of all reviews was $\acute{Z} = \{-6.994, 24.030, 20.237, -45.129, -65.474, -75.968, -162.301, -122.912, -24.824, -51.349, 100.327, -112.788, 100.327, -51.608, -198.068, -55.763, 25.065, -101.895, -123.110, -139.368\}$. The normalized priority of all reviews was $\acute{\Gamma} = \{0.640, 0.744, 0.732, 0.513, 0.444, 0.409, 0.120, 0.252, 0.581, 0.492, 1.000, 0.286, 1.000, 0.491, 0.000, 0.477, 0.748, 0.322, 0.251, 0.197\}$. Ranked in descending order, the helpfulness of the 20 reviews regardless of emotional intensity was $R_{11} \succcurlyeq R_{13} \succ R_{17} \succ R_2 \succ R_3 \succ R_1 \succ R_9 \succ R_4 \succ R_{10} \succ R_{14} \succ R_{16} \succ R_5 \succ R_6 \succ R_{18} \succ R_{12} \succ R_8 \succ R_{19} \succ R_{20} \succ R_7 \succ R_{15}$.

Table 12. Comparative analysis of Rank with ADV and Rank without ADV.

Rank	Rank with ADV	Rank without ADV
1	$\boldsymbol{R_{11}}$	$\boldsymbol{R_{11}}$
2	$\boldsymbol{R_{13}}$	$\boldsymbol{R_{13}}$
3	$\boldsymbol{R_{17}}$	$\boldsymbol{R_{17}}$
4	R_1	R_2
5	R_2	R_3

* Bold reviews indicate direct evaluation review on the film.

A comparative analysis of the rankings is shown in Table 12. Let us call the rank in which we take into account the emotional intensity indicator as "Rank with ADV" and the rank that does not take into account the emotional intensity indicator as "Rank without ADV". First, a comparison of "Rank with ADV" and "Rank without ADV" show that, the rank of the top three reviews is consistent, which shows that within a certain ranking range (the top three), considering emotional intensity did not affect on the ranking results. Second, of the top five reviews, four reviews remain unchanged. People are accustomed to reading multiple reviews to comprehensively judge the value of a product. Therefore, the emotional intensity index has little influence on the ranking result under the condition that the content presented is roughly the same. The above two points prove that the rank of helpfulness of reviews differs but to a lesser extent whether or not an indicator of emotional intensity is selected. This indicator can be considered depending on the specific business environment, or it can be weighted according to business requirements based on the two ranking results. For the comprehensiveness of the study, follow-up experiments will be carried out based on considering the index of emotional intensity.

Table 13. Statistical results of the top 5 reviews under different parameters.

Rank	$\delta = 0.1$	$\delta = 0.5$	$\delta = 1$
1	R_{11}	R_{11}	R_{11}
2	R_{13}	R_{13}	R_{13}
3	R_{17}	R_{17}	R_{17}
4	R_1	R_1	R_1
5	R_2	R_2	R_2
6	$\boldsymbol{R_{10}}$	$\boldsymbol{R_3}$	$\boldsymbol{R_3}$
7	$\boldsymbol{R_3}$	$\boldsymbol{R_{10}}$	$\boldsymbol{R_9}$

* Bold fonts are ratings that rank differently when sorting by different parameters.

D. PARAMETER SENSITIVITY ANALYSIS

When constructing the profit-loss priority matrix, we set the loss attenuation coefficient as 1. However, since SO-ILES TODIM is a decision-making method with a parameter, this parameter directly affects the calculation of the profit-loss priority value and ultimately affects the ranking of reviews. Therefore, a sensitivity analysis was carried out. We set $\delta = \{0.1, 0.5, 1\}$. The results are shown in Table 13.

For the convenience of analysis, we enumerate the first occurrence of different ranking positions under different parameters; that is, the sorting position of the 7th place. Table 13 shows that when the SO-ILES TODIM method takes different parameters, the review content and ranking of the first 5 of the 7 reviews are the same. The review content under the three parameters was not the same until the 7th review. These results indicate that different parameter values have little influence on order, which proves that the SO-ILES TODIM method is stable.

In the 6th review, the content of the review rendered by the order starts to differ due to the parameters. We select the 7th review for analysis since its ranking is different under three parameters. When $\delta = 0.1$, the 7th review is R_3: " Telling the whole truth in front of everyone, he said, the Consort Yu is the real Consort Yu, but the Ba Wang Xiang Yu is the fake one. (*in Chinese:* 他竟当面一语点破, 虞姬是真 虞姬, 霸王是假霸王。)". When $\delta = 0.5$, the 7th review is R_{10}: " My majesty is doomed to die; what's the meaning of my survival? (*in Chinese:* 君王意气尽, 贱妾何聊生。)". When $\delta = 1$, the 7th review is R_9: " The best movie in China.(*in Chinese:* 最优秀的中国电影。)". (*Background: Yu Ji is the name of a concubine, which in modern contexts means a wife, and Yu Ji is the wife's real name. We can simply understand that Yu Ji is a more specific reference.*)

δ represents the loss attenuation coefficient, and a smaller value indicates that the decision-maker is more concerned about the loss. In other words, they tend to prioritize more direct evaluation reviews of a film, which provide a more accurate understanding. However, with the increasing attenuation coefficient of losses, the extent of the decision-makers' avoidance of losses is reduced; that is, they are less concerned about the loss. In other words, the consumers are not as concerned about obtaining an accurate understanding of the film. As both R_3 and R_{10} are direct evaluations of the film, and R_3 is more specific than R_{10}, R_3 is screened out at the smallest parameter ($\delta = 0.1$), R_{10} is filtered at the bigger parameter ($\delta = 0.5$). R_9 applies to all films; therefore, it is an indirect review of a film, which means that we cannot judge the film from the reviews. Therefore, it is screened out when the decision maker is less concerned about the loss. The experimental results show that the loss attenuation coefficient can reflect the extent of decision-makers' loss avoidance, which indicates the scientific nature of the SO-ILES TODIM method.

V. CONCLUSION

Online reviews are an important basis for consumers to make purchasing decisions when shopping online. This paper studies the helpfulness ranking of online reviews to improve the purchasing efficiency by prioritizing helpful reviews. The research in this paper extends the research depth of the helpfulness of reviews, enriches the research method of the helpfulness ranking of reviews, and provides insights about the effective management of online reviews by businesses.

Taking film reviews as the research object, the SO-ILES TODIM method is proposed to rank the helpfulness of reviews. This method constructs a new language evaluation set, the intuitive language evaluation set based on emotional and ontological features (SO-ILES), which can effectively extract the characteristic information of research objects and is more applicable in the field. In addition, this method includes a calculation formula for index attribute value based on statistical rules and proposes the calculation method of index weight based on the logit regression model. These two points realize the quantitative calculation of attribute value and weight value, which, effectively avoids the subjectivity of manual assignment.

The case analysis demonstrates that the SO-ILES TODIM method can prioritize direct evaluation reviews of a film, which proves the effectiveness of the SO-ILES TODIM method. A comparative analysis of the choice of the emotional intensity index shows that its effect on the final review ranking is not significant. The choice of the emotional intensity index can be made according to the business environment. The parameter sensitivity analysis shows that the loss attenuation coefficient can not only ensure that the parameters reflect the decision maker's loss avoidance psychology but also ensure the relative stability of the review ordering in a certain range when the parameters change, which proves that the SO-ILES TODIM method is scientific.

There are still some deficiencies in the study. Due to the lack of review language standardization, a value of zero is prone to appear in the statistics of the evaluation index, thus causing the ranking equivalence problem. It is hoped that this problem can be improved in future studies. Additionally, due to people' s reading habits and the limitation of report space, the number of reviews selected in the study is relatively small. In the future, we can consider increasing the number of reviews to observe the consistency of conclusions.

REFERENCES

[1] K. Li, Y. Chen, and L. Zhang, "Exploring the influence of online reviews and motivating factors on sales: A meta-analytic study and the moderating role of product category," *J. Retailing Consum. Services*, vol. 55, Jul. 2020, Art. no. 102107.

[2] J.-W. Bi, Y. Liu, Z.-P. Fan, and E. Cambria, "Modelling customer satisfaction from online reviews using ensemble neural network and effect-based kano model," *Int. J. Prod. Res.*, vol. 57, no. 22, pp. 7068–7088, Nov. 2019.

[3] J.-W. Bi, Y. Liu, and Z.-P. Fan, "Representing sentiment analysis results of online reviews using interval type-2 fuzzy numbers and its application to product ranking," *Inf. Sci.*, vol. 504, pp. 293–307, Dec. 2019.

[4] E. Kauffmann, J. Peral, D. Gil, A. Ferrández, R. Sellers, and H. Mora, "A framework for big data analytics in commercial social networks: A case study on sentiment analysis and fake review detection for marketing decision-making," *Ind. Marketing Manage.*, vol. 90, pp. 523–537, Oct. 2020, doi: 10.1016/j.indmarman.2019.08.003.

[5] Y. Liu, J.-W. Bi, and Z.-P. Fan, "Ranking products through online reviews: A method based on sentiment analysis technique and intuitionistic fuzzy set theory," *Inf. Fusion*, vol. 36, pp. 149–161, Jul. 2017.

[6] Y. Liu, J. Bi, and Z. Fan, "A method for ranking products through online reviews based on sentiment classification and interval-valued intuitionistic fuzzy TOPSIS," *Int. J. Inf. Technol. Decis. Making*, vol. 16, no. 6, pp. 1497–1522, 2017.

[7] A. K. J and A. S, "Aspect-based opinion ranking framework for product reviews using a Spearman's rank correlation coefficient method," *Inf. Sci.*, vols. 460–461, pp. 23–41, Sep. 2018.

[8] C. Huang and W. Jiang, "Aspect-based personalized review ranking," in *Proc. IEEE SmartWorld, Ubiquitous Intell.Comput., Adv. Trusted Comput., Scalable Comput. Commun., Cloud Big Data Comput., Internet People Smart City Innov. (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Oct. 2018, pp. 1329–1334. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8560209

[9] S. Saumya, J. P. Singh, A. M. Baabdullah, N. P. Rana, and Y. K. Dwivedi, "Ranking online consumer reviews," *Electron. Commerce Res. Appl.*, vol. 29, pp. 78–89, May 2018.

[10] W. Wang, H. Wang, and Y. Song, "Ranking product aspects through sentiment analysis of online reviews," *J. Experim. Theor. Artif. Intell.*, vol. 29, no. 2, pp. 227–246, Mar. 2017.

[11] J. Singh, S. Irani, P. Rana, Y. Dwivedi, S. Saumya, and P. Roy, "Predicting the 'helpfulness' of online consumer reviews," *J. Bus. Res.*, vol. 70, pp. 346–355, Jan. 2017.

[12] Y. Shaalan, X. Zhang, and J. Chan, "Learning to rank items of minimal reviews using weak supervision," in *Advances in Knowledge Discovery and Data Mining*, vol. 10937. Cham, Switzerland: Springer, 2018, pp. 631–643, doi: 10.1007/978-3-319-93034-3_50.

[13] J. Bi, Y. Liu, and H. Li, "Daily tourism volume forecasting for tourist attractions," *Ann. Tourism Res.*, vol. 83, Jul. 2020, Art. no. 102923.

[14] M. Lin, C. Huang, and Z. Xu, "MULTIMOORA based MCDM model for site selection of car sharing station under picture fuzzy environment," *Sustain. Cities Soc.*, vol. 53, Feb. 2020, Art. no. 101873.

[15] J.-W. Bi, Y. Liu, Z.-P. Fan, and J. Zhang, "Exploring asymmetric effects of attribute performance on customer satisfaction in the hotel industry," *Tourism Manage.*, vol. 77, Apr. 2020, Art. no. 104006.

[16] M. Lin, C. Huang, Z. Xu, and R. Chen, "Evaluating IoT platforms using integrated probabilistic linguistic MCDM method," *IEEE Internet Things J.*, vol. 7, no. 11, pp. 11195–11208, Nov. 2020.

[17] J.-W. Bi, Y. Liu, Z.-P. Fan, and J. Zhang, "Wisdom of crowds: Conducting importance-performance analysis (IPA) through online reviews," *Tourism Manage.*, vol. 70, pp. 460–478, Feb. 2019.

[18] Q. Wu, W. Lin, L. Zhou, Y. Chen, and H. Chen, "Enhancing multiple attribute group decision making flexibility based on information fusion technique and hesitant pythagorean fuzzy sets," *Comput. Ind. Eng.*, vol. 127, pp. 954–970, Jan. 2019.

[19] M. Lin, Z. Chen, H. Liao, and Z. Xu, "ELECTRE II method to deal with probabilistic linguistic term sets and its application to edge computing," *Nonlinear Dyn.*, vol. 96, no. 3, pp. 2125–2143, May 2019.

[20] Q. Wu, L. Zhou, Y. Chen, and H. Chen, "An integrated approach to green supplier selection based on the interval type-2 fuzzy best-worst and extended VIKOR methods," *Inf. Sci.*, vol. 502, pp. 394–417, Oct. 2019.
[21] P. Liu and F. Teng, "Probabilistic linguistic TODIM method for selecting products through online product reviews," *Inf. Sci.*, vol. 485, pp. 441–455, Jun. 2019.
[22] D. Zhang, C. Wu, and J. Liu, "Ranking products with online reviews: A novel method based on hesitant fuzzy set and sentiment word framework," *J. Oper. Res. Soc.*, vol. 71, no. 3, pp. 528–542, Mar. 2020.
[23] C. Wu and D. Zhang, "Ranking products with IF-based sentiment word framework and TODIM method," *Kybernetes*, vol. 48, no. 5, pp. 990–1010, May 2019.
[24] M. Lin, Z. Xu, Y. Zhai, and Z. Yao, "Multi-attribute group decision-making under probabilistic uncertain linguistic environment," *J. Oper. Res. Soc.*, vol. 69, no. 2, pp. 157–170, Feb. 2018.
[25] R. Davoudabadi, S. Mousavi, and V. Mohagheghi, "A new last aggregation method of multi-attributes group decision making based on concepts of TODIM, WASPAS and TOPSIS under interval-valued intuitionistic fuzzy uncertainty," *Knowl. Inf. Syst.*, vol. 64, no. 4, pp. 1371–1391, 2020.
[26] J.-J. Wang, Z.-H. Miao, F.-B. Cui, and H.-C. Liu, "Robot evaluation and selection with entropy-based combination weighting and cloud TODIM approach," *Entropy*, vol. 20, no. 5, p. 349, May 2018.
[27] Y. Wu, J. Wang, S. Ji, Z. Song, and Y. Ke, "Optimal investment selection of industrial and commercial rooftop distributed PV project based on combination weights and cloud-TODIM model from SMEs' perspectives," *J. Cleaner Prod.*, vol. 234, pp. 534–548, Oct. 2019.
[28] P. Xiao, Q. Wu, H. Li, L. Zhou, Z. Tao, and J. Liu, "Novel hesitant fuzzy linguistic multi-attribute group decision making method based on improved supplementary regulation and operational laws," *IEEE Access*, vol. 7, pp. 32922–32940, 2019.
[29] D. Zindani, S. Maity, and S. Bhowmik, "Interval-valued intuitionistic fuzzy TODIM method based on Schweizer–Sklar power aggregation operators and their applications to group decision making," *Soft Comput.*, vol. 24, pp. 14091–14133, Mar. 2020.
[30] Z. Lu, Y. Gao, and W. Zhao, "A TODIM-based approach for environmental impact assessment of pumped hydro energy storage plant," *J. Cleaner Prod.*, vol. 248, Mar. 2020, Art. no. 119265.
[31] Y. Liang, Y. Tu, Y. Ju, and W. Shen, "A multi-granularity proportional hesitant fuzzy linguistic TODIM method and its application to emergency decision making," *Int. J. Disaster Risk Reduction*, vol. 36, May 2019, Art. no. 101081.
[32] C. Ruiz-Mafe, E. Bigne-Alcaniz, and R. Curras-Perez, "The effect of emotions, eWOM quality, and online review sequence on consumer intention to follow advice obtained from digital services," *J. Service Manage.*, vol. 31, no. 3, pp. 465–487, 2020, doi: 10.1108/JOSM-11-2018-0349.
[33] Y. Yang, C. Chen, and F. S. Bao, "Aspect-based helpfulness prediction for online product reviews," in *Proc. IEEE 28th Int. Conf. Tools with Artif. Intell. (ICTAI)*, Nov. 2016, pp. 836–843. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/7814690
[34] J. M. Kim, K. Hwang, and S. Hyun, "Understanding the dynamics of review posting and the management intervention effects on those dynamics," *Int. J. Hospitality Manage.*, vol. 88, Jul. 2020, Art. no. 102526.
[35] A. Han, R. Jifan, and Z. Mengyang, "The impact factors of online book reviews usefulness: An empirical comparison between ANN and SVM," in *Proc. 15th Int. Conf. Service Syst. Service Manage. (ICSSSM)*, Jul. 2018, pp. 1–6. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/8464990
[36] X. Sun, M. Han, and J. Feng, "Helpfulness of online reviews: Examining review informativeness and classification thresholds by search products and experience products," *Decis. Support Syst.*, vol. 124, Sep. 2019, Art. no. 113099.
[37] Y. Zhou and S. Yang, "Roles of review numerical and textual characteristics on review helpfulness across three different types of reviews," *IEEE Access*, vol. 7, pp. 27769–27780, 2019.
[38] Y. Hou, H. Dong, M. Hao, and X. Cui, "Evolutionary stable strategies and game dynamics," *Res. Fine-grained Emotion Classification Film Rev. Based Ontol. Features*, vol. 40, no. 4, pp. 1074–1078, 2020.
[39] D. Zhang, Y. Li, and C. Wu, "An extended TODIM method to rank products with online reviews under intuitionistic fuzzy environment," *J. Oper. Res. Soc.*, vol. 71, no. 2, pp. 322–334, Feb. 2020.
[40] M. Lin, H. Wang, and Z. Xu, "TODIM-based multi-criteria decision-making method with hesitant fuzzy linguistic term sets," *Artif. Intell. Rev.*, vol. 53, no. 5, pp. 3647–3671, Jun. 2020.

HUIFANG DONG received the master's degree in management science and engineering from the Shandong University of Science and Technology. Currently, she has published one academic article in *Journal of Computer Applications*. Her research interest includes text analysis of user-generated content.

YANHUI HOU is currently an Associate Professor and a Master Supervisor with the Shandong University of Science and Technology, Qingdao, China. He had presided over or participated in nearly ten projects of the Natural Science Foundation of China and the Province Science Foundation of Shandong, and more than 20 projects entrusted by enterprises. He has published more than 20 academic articles and one academic monograph (chief editor). He has been involved in the research and teaching of management science and engineering for a long time. His research interests include the behavior law of management systems and system engineering. He has won the China Coal Industry science and technology Progress Award and other awards for five times.

MIN HAO is currently pursuing the Ph.D. degree in management science and engineering with the Shandong University of Science and Technology. She is currently a Lecturer with the Shandong University of Science and Technology. She has participated in more than ten scientific research and teaching research projects, and nearly ten projects entrusted by enterprises. She has published more than 20 academic articles, published one work (the Deputy Chief Editor), and participated in the compilation of one teaching material. Her research interests include systematic evaluation methods and applications, and text mining.

JIAKUN WANG received the Ph.D. degree in management science and engineering from the Shandong University of Science and Technology. He has published more than ten academic articles in the international journal so far, such as *Physica A: Statistical Mechanics and Its Applications*, *Safety Science*, and so on. His research interests include complex networks, spreading dynamics, and public opinion propagation and management.

SHUOSHUO LI received the master's degree in management science and engineering from the Shandong University of Science and Technology. He has published many academic articles in *China Population, Resources and Environment* and other domestic journals. His research interests include green innovation and collaborative innovation networks, innovation ecosystem networks, and evolution.