Explanation Chains Model based on the Fine-grained Data[1]

Ma Fu-Yuan[1], Chen Wen-Qi[1], Xiao Min-Hao[1], Wang Xin[2,3] and Wang Ying[1,2,4]

[1] College of Computer Science and Technology, Jilin University, Changchun 130012
[2] Key Laboratory of Symbol Computation and Knowledge Engineering（Jilin University）, Ministry of Education, Changchun 130012
[3] Changchun Institute of Technology, Changchun130012
[4] College of Software, Jilin University, Changchun 130012
wangying2010@jlu.edu.cn

Abstract. With the development of information society, Recommendation System has been an import tool to help users filter information and create more economic value for enterprises. However, it is difficult for traditional recommendation systems to interpret recommendation results. In order to improve users' trust in recommendation results, interpretable recommendation models have attracted more and more attention. In this paper, we present the Explanation Chains Model based on the Fine-grained Data (F-ECM) to enhance the effectiveness of recommendation while achieving the interpretability of recommendation. First, we generate parsing trees from user comments and extract three key sentence structure information (i.e, aspects, features and sentiment tendency) from those generated parsing tree. The fine-grained similarity is computed based on the aspects and features of the products to be recommended， and users' product satisfaction is predicted by combining sentiment tendency. Then the recommendation chain will be constructed according to the satisfaction degree in the recommendation list. Finally, we calculate recommendation chain scores of all the items to be recommended to the target user, generate the recommendation results and personal explanation for the user by the recommendation chains. Experiments in the Amazon data set show that the Explanation Chains Model based on the Fine-grained Data achieve better interpretability and performance of product recommendation systems.

Keywords: Recommendation System; Interpretation; Explanation Chain; Fine-grained Data Set

1 Introduction

With the advent of the information age, the amount of information on the Internet has increased explosively [1]. In order to deal with the problem of "information overload" and filter data quickly and effectively, recommendation system has been widely used as an effective method [2].Currently, existing recommendation systems focus on end-to-end optimization using evaluation indicators. For example, root mean square error is used to predict user ratings of products and differences between real ratings. However, user's decision is related to many factors. End-to-end recommendation model only implicitly utilizes these factors, which can't explain explicitly which factors have an impact on recommendation results. Although many algorithms have improved the performance of indicators [3] [4] [5], when the algorithm produces poor recommendation results, it will directly lead to a decline in user acceptance of the recommendation system. In order to improve the persuasiveness of recommendation, the algorithm should explain to users why recommendation occurs, even if deviation does not unduly affect users' trust.

User's comments contain abundant information, such as product description and customer sentiment. These fine-grained data at the product level are of great significance for both of user analysis and product evaluation. A summarization system based on user evaluation can perceive the descriptors of the product that user is interested in and the satisfaction degree with the descriptors of the product and summarize all the comments on a product. In this way, it can perceive the various aspects of a particular product descriptor and its corresponding popularity. Therefore, we use fine-grained data to construct user and product descriptors, and our recommendation is based on users' sentiment tendency. The recommendation effect may be improved by utilizing the popularity of the various aspects of a particular product descriptor so as to enhance the interpretation of the recommendation system. And the recommendation system is able to recommend the products based the similarity of the descriptor aspects, meanwhile it may provide a more reasonable and convincing recommendation results. We proposes an Explanation Chains Model based on the Fine-grained Data to the interpretability of the recommendation system. The main contributions are threefold:

1. We extract the syntactic structure rules and use an unified structure to represent the user comments.
2. We propose a method for extracting key aspects from the unified structure representation and obtain fine-grained data of the user's comments on product.

[1] This work was supported by the National Natural Science Foundation of China (61872161, 61602057), the Science and Technology Development Plan Project of Jilin Province (2018101328JC), the Science and Technology Department Excellent Youth Talent Foundation of Jilin Province (20170520059JH), the Project of Technical Tackle-Key-Problem of Jilin Province (20190302029GX), and the Project of Development and Reform of Jilin Province (2019C053-8).

3. We propose the F-ECM, and make full use of the multiple dimensions of fine-grained data to recommend, which not only improves the recommendation effect, but also provides users with readable product characteristics of the recommendation reason.

Section 2 presents related work, Section 3 is description of the problem, Section 4 is the Explanation Chains Model based on the Fine-grained Data.

2 Related Work

2.1 Interpretability of Recommendation System

With the in-depth study of recommendation system, the validity and accuracy of recommendation results are gradually improved. However, the traditional recommendation system does not provide a reasonable explanation of the recommendation results, which affects the user's acceptance of the recommendation results.

The research on the transparency of the recommendation system [6] and the explanatory research on the commonly used collaborative filtering algorithm [7] have achieved certain results. The early recommendation system provides explanatory by using the product's content tag, it can be used both for recommendation prediction and for explaining to the user [8]. Zhang et al. [9] formally defined interpretability in 2014: explaining how the system works and/or why a product is recommended, the system becomes more transparent and has the potential to allow users to tell when the system is wrong (scrutability), increase users' confidence or trust in the system, help users make better (effectiveness) and faster (efficiency) decisions, convince users to try or buy (persuasiveness), or increase the ease of the user enjoyment (satisfaction). In recent years, the explainable of recommendation systems has been systematically classified and summarized [10]. Explainable recommendation system can be classified into matrix factorization, topic modeling, graph-based, and deep learning, etc.

2.2 Explainable Recommendation Model

The topic modeling-based explainable recommendation model mainly obtains recommendation results and explanations by analyzing and combining the content topics of the text. McAuley and Leskovec [11] proposed a method to obtain the recommended interpretation by aligning the implicit aspect descriptors of the recommended decomposition with the implicit topics of the Latent Dirichlet allocation (LDA). Lin Li et al. [12] proposed DTMF + model, which established a positive mapping relationship between the potential topic vectors of the user comment set and the product review set and the user potential factor vector and product potential factor vector of the traditional matrix factorization, respectively, and further guided scoring prediction by adding potential topics. Tan et al. [13] proposed the concept of "recommendability" of products. Based on boosting method, the characteristic distribution of the product recommendability is linked to the user's preference distribution in the same space. At the same time, users' ratings and comments are used for collaborative recommendation.

There are connections between users and users, products and products, users and products. It is easy to construct a graph structure, especially in the social recommendation task, using the graph structure to make the user and product modeling more intuitive. He et al. [14] obtained the Top@K recommendation model by introducing the ternary relationship of the three-dimensional graph modeling (user-product-side). Wang et al. [15] proposed a tree-enhanced embedded model to obtain explainable recommendations, combining the generalization ability of the embedded model with the interpretability of the tree model. The Explanation Chains algorithm proposed by Arpit [16] is an interpretation-based recommendation algorithm that calculates recommendations by generating and sorting explanation chains which are generated by modeling the recommendation problem as a question of finding a path unifies recommendations and explanations in a productsimilarity chart.

Matrix factorization algorithm and its variants have achieved great success in recommendation tasks, including classical latent factor model LFM [17], singular value decomposition model (SVD) [18][19], non-negative matrix factorization model NMF[20]. The recommended prediction results of these models have higher accuracy, but the factor obtained by the factorization is an implicit description. It is difficult to explain the reasons for the recommendation to the user, and the user needs a certain degree of trust in the recommendation system.

In recent years, deep learning and presentation learning have received widespread attention, and an explainable recommendation system using deep learning techniques has gradually emerged. Combined with CNN[21], RNN[22], attention mechanism[23], etc., a variety of recommendation interpretation modes are generated, for example, automatic generation of interpretation text without the template[24], visualization of interpretation based on user's attention on images without interpretation template[25], etc.

3 Problem Describe

Given a user comment data set D, first, extracting one or more sets of fine-grained triples $(a, f, s)_{ij}$ from every comment text in D which indicates that user i uses descriptor f for aspect a of item j and showing an emotional tendency s. Then, using the fine-grained data triple of the user's comment on the product, the user is represented by the evaluation content of the product evaluated by each user to establish the user data U, wherein, for the i-th user, the number set of the product evaluated by the user is represented by item-i, then $U_i = \{(a, f, s) \mid (a, f, s)_{ij}, j \in \text{item-i}\}$; the product is represented by the specific content received from a plurality of users of the product, wherein, for the j-th product, the number set of the user who has evaluated the product is represented by user-j, then $I_j = \{(a, f, s) \mid (a, f, s)_{ij}, i \in \text{user-j}\}$. Next, U and I are used to establish a recommendation chain C_{ik} for the product k that the user i has not touched, and the product k is swam in the product evaluated by the user i, meanwhile, the recommendation chain is established by using the predicted satisfaction (ps). The recommendation chain is a list. Its head is the product to be recommended, and the other items are composed of the product evaluation content in U_i. Finally, the recommended chain is evaluated to produce the recommended product and the recommended explanation.

Table 1. Symbol Table.

Symbol	Meaning
D	User-to-product review data set
(a, f, s)	Fine-grained reviews of triples (aspects, aspect descriptors, sentimental tendencies)
f_f	Aspect frequency of occurrence
U	User data set
item-i	Number set of the product evaluated by the user i
U_i	Specific composition of user i in U：$U_i = \{(a, f, s) \mid (a, f, s) \in \text{review}_{ij}, j \in \text{item-r-by-i}\}$
I	product data set
user-j	Number set of the user who has evaluated the item j
I_j	Specific composition of item j in I：$I_j = \{(a, f, s) \mid (a, f, s) \in \text{review}_{ij}, i \in \text{user-r-j}\}$
C_{ik}	Recommendation chain built for product k that user i has not touched
ps(j,k)	The predicted satisfaction of the product k calculated based on the product j evaluated by the target user

4 Explanation Chains Model based on the Fine-grained Data

By processing the user text to obtain the product description words and corresponding sentiment trends discussed by the user in the comments, a fine-grained data set is established. The recommendation chain model is improved by making full use of various dimensions of the fine-grained data, so it provides a more informative and accurate recommendation for recommended products.

The model we built is shown in Figure 1：

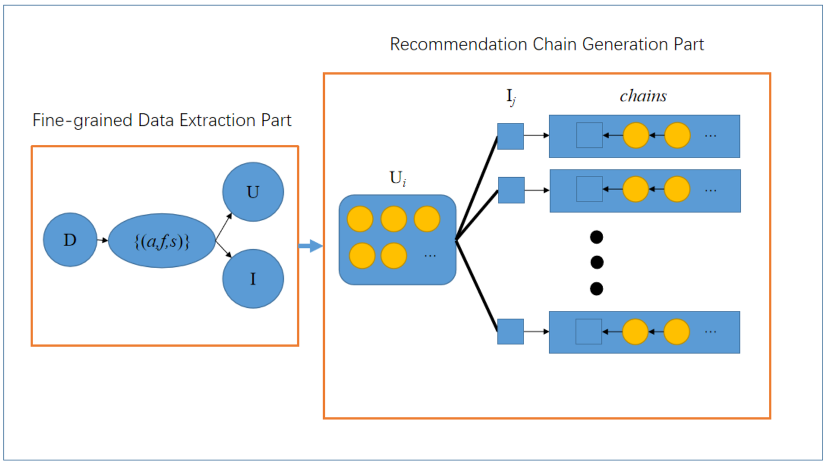

Fig. 1. Explanation Chains Model based on the Fine-grained Data

4.1 Fine-grained data Extraction

The quality of fine-grained data directly affects the effect of the algorithm. We extract fine-grained data based on the principle of retaining as much information as possible. Although the expression of natural language is very flexible, some euphemistic expressions or discourses in special situations are very difficult to conduct structured analysis, so it is difficult for the machine to obtain accurate user expression. But in the specific context of product reviews, the user's expression often has a good structure, which is convenient for research and analysis.

As shown in Figure 1, the fine-grained data extraction part is mainly completed in the following steps.

Step1 Text preprocessing to remove irrelevant expressions in text;

Step2 Syntactic structure parsing, using the parse tree to represent the comment text;

Step3 Extracting key information in the sentence by the common structural rules in the parsing tree to obtain a unified structured representation of comment text;

Step4 Extracting a three-tuple set (aspect, aspect descriptor, sentiment tendencies) from a unified structured representation of the review text.

Step1 standardizes the comment text, removes the non-utf-8 encoded characters that may exist in the text, converts the ordinary text into a uniform lowercase, restores the irregular word spelling and restores the English abbreviations. Step2 uses Stanford University's English parsing tool [26] to parse the comment text and get the parsing tree.

Step3 regards aspects of goods, aspect descriptors and emotional tendencies as key information of fine-grained data. Aspect is a property of a product, and often presented as a noun or pronoun in a sentence. Aspect descriptors are supplements to a description of a particular aspect of a product, usually appears around the aspect words as adjectives or adverbs for modification. The user's emotional tendency has two sources, one is the aspect descriptors that the user used to describe the aspect of the product, and the second is the verb used in the user description, such as: "like", "hate", "repent" and other words. In the process of extraction, in order to preserve the key information of the text as much as possible, we select a certain number of example sentences to analyze the relevant conclusions about the sentence composition in the parse tree, and proves the conclusion by statistical analysis of the data set. These conclusions propose ways to unify the expression of the text. The conclusions reached are as follows:

Conclusion 1: A sentence consists of two main components: noun phrase (NP) and verb phrase (VP)

Conclusion 2: Noun phrases are mainly composed of the following part of speech: pronoun / noun / qualifier / adjective

Conclusion 3: The verb phrase is mainly composed of the following components: verb/noun phrase/preposition phrase/adverb phrase/sentence/adjective phrase

Based on the above conclusions, we provide a unified and standardized representation of the review texts in order to extract the aspects, features and emotional tendency of the product mentioned by user. According to the conclusion 1, a sentence has two main parts — the noun phrase and the verb phrase. Therefore, according to the expression habit, we recognize the two as the subject and the predicate, and the main structure is obtained: [subject noun phrase], [predicate verb phrase].

According to the conclusion 2, the composition of the noun phrase part is formatted, the three main parts of the noun phrase are pronoun, noun, adjective. Pronouns and nouns represent real things, so they are treated as the same nature and placed in the same position. The expression nature of adjectives is different from nouns and it is of great significance to the expression of features in various aspects. So we store it together with nouns in noun phrases. Therefore, the structure of the final noun phrase is reserved as [adjective, noun/pronoun].

According to the conclusion 3, there are two kinds of different words in the verb phrase, one is the verb that expresses the user action and even the emotion, the other is the simplification and retention of the noun phrase as the action object. Verbs retain themselves, meanwhile, considering adverbs and adjectives that have a high proportion in verb phrases and play an important role in emotional expression, we finally decide to retain the verb phrase components as [[adverb, verb], adjective] format. The noun phrase appearing in the verb phrase is the same as the form of conclusion 2 as [adjective, noun/pronoun].

Finally, the sentence is uniformly normalized as the following structure: {[adjective, subject (noun/pronoun)], [[adverb, predicate (verb), adjective], [adjective, object (noun/pronoun)]}, in the sentence the components of the aspect/aspect descriptor/emotional expression potential are retained as much as possible. For the compound sentence structure, it is divided into multiple simple sentences and saved one by one, which does not affect the components of the sentence that express the key information.

Step 4 extracts all the (aspects, aspects, and emotions) triples in the structure from the results of Step 3 for subsequent use and extraction of key information. First, the core of a unified normalized structure is the predicate part, which connects the action subject and the object. First, the predicate part in the unified normalized structure is retained as the possible emotional expression in the emotional part of the second unified normalized structure. Then, extract all existing aspect descriptors--the aspect from the residual structure, that is, the adjective-noun combination, and ignore the content in the subject or object when it is a personal pronoun. If it is another pronoun, use the category of the product instead. At this time, if the adjective part is not empty, the two are combined into one aspect and aspect description, otherwise it is skipped; if the noun part is a specific noun, it is directly retained as a product, and if the adjective part is not empty, retain the description words for this aspect. For an adverb or adjective around a predicate verb, if it is not empty, it is combined with the noun in the subject to produce an aspect—aspect descriptor combination. If the noun is empty, the product category is used as a noun part. Finally, for a sentence, the predicate is combined with all aspect-side descriptors to form a triplet (verb, adjective, noun).

4.2 Recommendation Chain Generation Combined with Fine-grained Data

As shown in the recommendation chain generation part in Fig. 1, given a target user U_i, a project to be recommended I_j Firstly, F-ECM will build a recommendation chain from the project, and then calculate the target user's prediction satisfaction with the item based on the product evaluated by the target user. And select the target user's evaluated product to make the project's prediction satisfaction highest, if the predicted satisfaction calculated by the project is greater than the set threshold, then it will be added to the recommendation chain, saying that the newly added item is a precursor item and is represented by the product data set I; otherwise, the recommendation chain is established. Then, according to the products that the target user has evaluated and not in the current recommendation chain, the user's predicted satisfaction with the predecessor items is calculated in turn, and the item with the highest satisfaction degree is selected. If the predicted satisfaction is greater than the threshold, the recommendation chain is added and the newly added project is regarded as a precursor project, otherwise the recommendation chain is established; the above process is repeated until the recommendation chain is established.

As shown in Figure 2, the target user U_i is represented by the product it evaluated. When establishing the recommendation chain for I_j, the target user's predictive satisfaction with I_j is calculated according to I_k, I_l and other products representing U_i (which are represented by the triple representation extracted from the target user's evaluation text). After calculating, I_k in Figure 2 which has the highest predictive satisfaction is selected and added to the recommendation chain and I_k as the precursor product. At this time, I_k as the precursor product is represented by the product data set, i.e, the triple representation in the evaluation text of the product by all users. Then, according to the items that constitute U_i and are not in the chain, the predicted satisfaction of the current precursor product I_k is calculated in turn according to the items that constitute U_i and are not in the chain. After the calculation is completed, the I_l in Figure 2, the product with the highest satisfaction, is selected to continue the above process as a new precursor product. If all the products representing users are in the chain or the highest satisfaction calculated for the precursor product fails to reach the threshold value, stop the process and complete the recommendation chain.

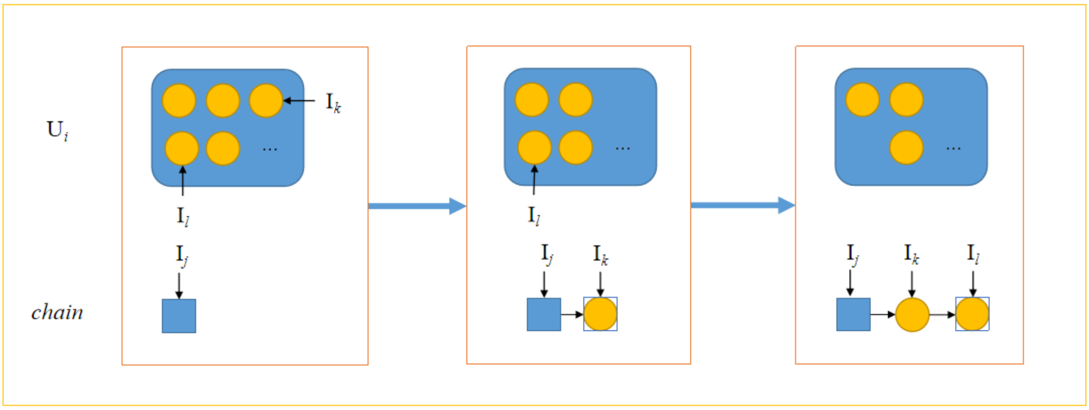

Fig. 2. Establishment of the Recommendation Chain

When building recommendation chain, we use similarity between goods as a criterion to select items added to the chain. Its intention is to use similarity between goods to provide an explanation for recommendation. In the fine-grained data that we used, the user's emotional inclination to aspects is mined, which can generate more accurate recommendation interpretation for users at the user's preference level. User's previous evaluation records refer to the products that the user cares about. All evaluations received by a product in some way can be used as a word of mouth for a product in this respect. When recommending a product to users, users can be predicted by analyzing the specific performance of the product to be recommended in terms of users' likes or dislikes. To quantify the satisfaction degree of products is to predict the satisfaction degree, based on fine-grained data, we propose ps_{ij} to predict the satisfaction degree of recommended product *i* according to user's historical product *j*.

In order to calculate ps_{ij}, firstly, considering the similarity of two products in aspect, then calculating the similarity between aspect descriptors based on aspect similarity. Finally, combining the user's emotional tendency towards aspect descriptors and the word-of-mouth of products on aspect descriptors, the prediction satisfaction is calculated. The specific methods are as follows:

There is a one-to-many relationship between aspects and aspect descriptors in fine-grained data. We organize product aspects, aspect descriptors, and emotional tendencies in the following ways:

$$\{a_1{:}\{[f_1, f_{f_1}, s_1],\ [f_2, f_{f_2}, s_2],\ \ldots\},\ a_2{:}\{\ldots\}, \ldots\}$$

That is, the product aspect is used as a first-level element describing the product, and in each product aspect, there are a plurality of aspect descriptors for the aspect as the secondary element.

Therefore, when calculating the predicted satisfaction of the recommended product based on the product (historical product) evaluated by a target user, first, calculating the similarity of all aspects of the predicted product and the user's reference in the historical product, retaining the combination of aspects the similarity of which beyond threshold . Then, for each of the remaining combinations of merchandise, calculating the similarity between the descriptors of all aspects in these two aspects. When calculating, consider the Frequency of occurrence of each historical product to indicate the user's Attention level of the descriptor, the frequency of the aspect descriptors on the aspect of each recommended item indicates the degree of recognition of the descriptor. Therefore, the proportion of the frequency of the description words in the corresponding aspects of the historical product indicates that the user cares about the description of the aspect, the frequency of occurrence of the descriptors of the recommended items indicates the degree of recognition of the descriptors, and the above factors are used to obtain the similarity of the two products in terms of aspects. Finally, combining l sentimental tendency and word-of-mouth to predict satisfaction, this paper divides sentimental tendency into three levels: positive, neutral, and negative, and uses 1, 0, -1 to express them respectively. While calculating, considering the following: If there are two items whose sentimental tendency of the aspect descriptor are consistent in a similar aspect, then the user's possible satisfaction of this similar aspect is consistent with the tendency of descriptors in this aspect.

If the emotional tendency of the two items in a similar aspect is inconsistent, then using the difference between the comprehensive sentimental tendency of the product to be recommended and the emotional tendency of the descriptor of the historical product to indicate the user's possible satisfaction, if the emotional expression of the historical product is negative, and the overall emotional sentiment of the product to be recommended is positive. In this respect, the good-to-recommended goods will be more attractive. Conversely, the descriptor of this aspect is a positive tendency, and the overall sentimental tendency of the product to be recommended is a negative tendency, then the aspect descriptor which is not good in this aspect will have more negative effects on the user who refers to the descriptor. Therefore, this paper uses the difference between the comprehensive sentimental tendency of the product to be recommended and the sentimental tendency of the descriptor in the historical product to achieve the emotion-related calculation in the forecasting satisfaction.

Finally, the user's historical product is defined as the predicted satisfaction of the recommended product $ps(j,i)$ for:

$$\mathrm{ps}(j,i)=\sum_{a_i\in item_i}\sum_{a_j\in item_j}(sim(a_i,a_j)*\sum_{f_m\in a_i}\sum_{f_n\in a_j}(sim(f_m,f_n)*sentidifi(f_m,f_n)*w_{f_n}*f_{f_m}))\quad(1)$$

Among them, $item_i$ represents the product evaluated by the users, $item_j$ represents the product to be recommended, ai is the aspect of $item_i$, a_j is the aspect of $item_j$, f_m is the descriptor of a_i , f_n is a descriptor of a_j, w_{f_n} is the proportion of the descriptor of the item aspect f_n in the descriptor of aspect a_j , f_{f_m} the frequency of occurrence of the user-side descriptor f_m in the descriptor of aspect a_i, and sim(a_i, a_j) represents the similarity between aspects a_i and a_j , sim(f_m, fn) indicates the similarity between the aspects f_m and f_n, and the computing method of *sentidifi*（f_1，f_2）is:

$$sentifi(f_1,f_2)=\begin{cases}senti(f_2)-senti(f_1) & if\quad senti(f_1)!=senti(f_2)\\ senti(f_1) & if\quad senti(f_1)=senti(f_2)\end{cases}\quad(2)$$

Among them, $senti(f_1)$ represents the sentimental of the user in the aspect descriptor f_1.

4.3 Recommendation Chain Generation Combined with Fine-grained Data

The assessment recommendation chain adopts the same method as the assessment interpretation chain. After establishing the recommendation chain for all the items to be recommended for the target users, using the non-chain products in the chain to calculate the predicted satisfaction of the products to be recommended，and calculate the average to express the score of recommendation chain ,select the n recommended chains with higher scores for top-n recommendation. Define score（<C,i>,C*）to describe and calculate the score of the recommendation chain with i as the chain header, where i is the item to be recommended, C is the recommendation chain with i as the chain header, and C* is the selected recommendation chain:

$$score(<C,i>,C^*)=\frac{\sum_{j\in C}ps(j,i)}{|C|+1}+\frac{C\setminus\cup\bigcup_{j'\in C^*}j'}{|C|+1}\quad(3)$$

Among them: the previous part represents the mean of the values of *ps* and *i* among all items in the chain, and the latter part of the pnalty chain has appeared in the recommendation chain in the final recommendation chain, while making the final recommended results cover the user's historical information as much as possible. The latter part diversifies the results while also reducing the impact of popular paranoia by not considering the items covered in the chain that have been selected in the formula.

5 Experiments and Results Analysis

5.1 Experimental Data

The experimental data in this article uses the Amazon product data: Cell Phones and Accessories. The final data set is summarized as follows:

Table 2. Data set properties

properties	number
Comment	12531
User	1117
Product	1232
Aspect	14573
Aspect Descriptor	8452
Simplified Expression Triplet	220211

Note: Simplified expression of the triplet is the final result of the comment processing: (verb, adjective, noun)

As can be seen from Table 2, a total of 220,211 triples were extracted from 12,531 user comments, with an average of 17 triples in each comment. It indicates that the method of extracting key information has kept the content of the user's comment text as much as possible.

5.2 Experimental Data

The example sentence is processed according to the method used in this paper to obtain a fine-grained data set:
Example: These are awesome and make my phone look so stylish!
Parsing Tree：

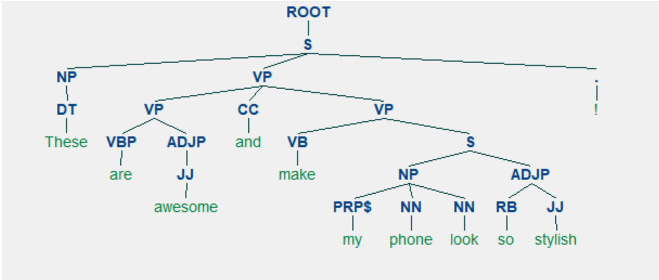

Fig. 3. An example of a parse tree

From the results of parsing, a unified normalized structure is obtained：

[[(' ', 'battery charger cases'), [[' ', 'are'], 'awesome'], (' ', ' ')], [(' ', 'battery charger cases'), [[' ', 'make'], 'stylish'], ('stylish', 'phone look')]]

Extract brief information from the unified normalized structure：

[('make', 'stylish', 'phone look'), ('are', 'awesome', 'battery charger cases'), ('make', 'stylish', 'battery charger cases')]

It can be seen from the above example that the simplified structure preserved after parsing is relatively complete for sentence information preservation, and the simplification and extraction effect of key information based on part of speech is also reasonable.

5.3 Experimental Data

Table 3 is the experimental results of using the MAE and RMSE indicators for 5-fold cross-validation on the Amazon data set. Comparison algorithm includes KNNBasic, KNNWithMeans, NMF, biasSVD, SVD++ and NormalPredictor. The F-ECM model has the best value on the MAE index, indicating that it has a good recommendation effect, but the performance on the RMSE indicator is general, indicating a large fluctuation in the recommended effect. The main reason is that F-ECM relies heavily on users' comments, while users' comments are arbitrary. The amount of content involved will directly affect the effect of recommendation to users. The richer the content of comments, the more accurate the recommendation to users will be.

Table 3. Comparison of Recommendation Algorithms

MAE					5-fold cross-validation on the Amazon data set			RMSE				
models	1	2	3	4	5	AVG	1	2	3	4	5	AVG
KNNBasic	0.6397	0.6352	0.6451	0.6484	0.6343	**0.6405**	0.9496	0.9272	0.9523	0.9437	0.9339	**0.9413**
KNNWithMeans	0.6545	0.6491	0.6753	0.6285	0.6476	**0.6510**	0.9637	0.9514	0.9882	0.9281	0.9578	**0.9578**
NMF	0.7412	0.7625	0.7373	0.7393	0.7394	**0.7439**	0.9951	1.0204	0.9969	0.9913	0.9949	**0.9997**
biasSVD	0.6293	0.6232	0.6170	0.6201	0.6136	**0.6206**	0.8831	0.8569	0.8565	0.8696	0.8369	**0.8606**
SVD++	0.6059	0.6183	0.6176	0.6167	0.6011	**0.6119**	0.8475	0.8671	0.8792	0.8598	0.8459	**0.8599**
NormalPredictor	0.9488	0.9547	0.9592	0.9322	0.9472	**0.9484**	1.2674	1.2787	1.2720	1.2488	1.2630	**1.2660**
F-ECM	0.5402	0.5499	0.3646	0.4271	0.4756	**0.4915**	1.0384	1.0558	0.8593	0.6578	1.0861	**0.9395**

5.4 Experimental Data

F-ECM can generate fine-grained recommendation reasons with user characteristics for any recommended product in the data set. It can be used as a supplementary model of interpretability in recommendation system.

For example, for users whose ID is "A2P68VRKQMYBDE":

When recommending the product with the product number "B00KGU9UHS", the effect of using radar image to show users intuitively is as follows：

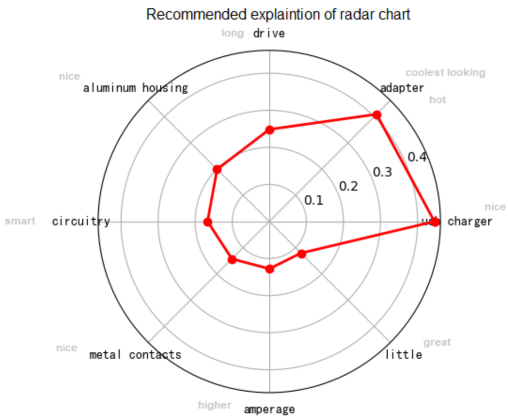

Fig. 4. Example of F-ECM

It can be seen from Figure 4 that the item to be recommended – charger, Most of the reasons for recommendation are highly relevant to the product: usb charger, adapter, drive, aluminum housing, circuitry, metal contacts, amperage. F-ECM provides users with a reference based on the aspect of the product, the aspect descriptor combined with the performance of the product in this aspect, the user's past preferences and the frequencies mentioned in all comments.

6 Conclusion

Through the natural language analysis of user comment texts, this paper obtains the fine-grained data set from the user's text expression, and realizes the interpretability of user recommendation by establishing the F-ECM. This model not only makes it easy for users to produce interpretable recommendations for goods, but can also be used as an explanatory complement to other recommendation systems.

Since the F-ECM mainly emphasizes the interpretable effect, there is a large demand for the users in the data set, and there are some limitations in the application process. The next research will mainly consider synthesizing fine-grained data and other data to produce interpretable recommendation results, and improve the generality and recommendation efficiency of interpretable models.

References

1. Marz N, Warren J. Big Data: “Principles and best practices of scalable realtime data systems. Greenwich,” USA: Manning Publications Co, 2015
2. Adomavicius G, Tuzhilin A. “Toward the next generation ofrecommender systems: A survey of the state- of-the-art and possible extensions.” IEEE Transactions on Knowledge and Data Engineering, 2005, 17(6), pp. 734-749
3. Sangeeta and N. Duhan, “Collaborative Filtering-Based Recommender System,” in Advances in Intelligent Systems and Computing, vol. 653, Berlin, Heidelberg: Springer Berlin Heidelberg, 2018, pp. 195–202.
4. B. Sarwar et al., “Item-Based Collaborative Filtering Recommendation,” WWW ’01 Proc. 10th Int. Conf. World Wide Web, 2001, pp. 285–295.
5. L. Baltrunas, B. Ludwig, and F. Ricci, “Matrix factorization techniques for context aware recommendation,” in Proceedings of the fifth ACM conference on Recommender systems - RecSys ’11, 2011, pp. 301.
6. K. Koizumi et al., “The role of presenilin 1 during somite segmentation.”, Development, Apr. 2001, vol. 128, no. 8, pp. 1391–402.
7. J. L. Herlocker, J. A. Konstan, and J. Riedl, “Explaining collaborative filtering recommendations,” in Proceedings of the 2000 ACM conference on Computer supported cooperative work - CSCW ’00, 2000, pp. 241–250.
8. B. Ferwerda, K. Swelsen, and E. Yang, “Explaining Content-Based Recommendations,” New York, pp. 1–24.
9. Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, and S. Ma, “Explicit factor models for explainable recommendation based on phrase-level sentiment analysis”, in Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval - SIGIR ’14, 2014, pp. 83–92.
10. Y. Zhang and X. Chen, “Explainable Recommendation: A Survey and New Perspectives” , CoRR abs/1804.11192, Apr.2018.
11. J. McAuley and J. Leskovec, “Hidden factors and hidden topics: understanding rating dimensions with review text” in Proceedings of the 7th ACM conference on Recommender systems - RecSys ’13, 2013, pp. 165–172.

12. L. Lin, L. Jin-Hang, M. Xiang-Fu, et al. “Recommendation Models by Exploiting Rating Matrix and Review Text.”Chinese Journal of Computers, 2018,v.41,No.427(07):pp.131-145. (in Chinese)
13. Y. Tan, M. Zhang, Y. Liu, and S. Ma, “Rating-boosted latent topics: Understanding users and items with ratings and reviews,” in IJCAI International Joint Conference on Artificial Intelligence, 2016, vol. 2016–Janua, pp. 2640–2646.
14. He X, Chen T , Kan M Y , et al. “TriRank: Review-aware Explainable Recommendation by Modeling Aspects[C]” the 24th ACM International. ACM, 2015.
15. X. Wang, X. He, F. Feng, L. Nie, and T.-S. Chua, “TEM：Tree-enhanced Embedding Model for Explainable Recommendation,” in Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW ’18, 2018, pp. 1543–1552.
16. Arpit Rana and Derek Bridge. 2017. “Explanation Chains: Recommendation by Explanation.” RecSys’17 Poster Proceedings, Como, Italy, August 27–31, 2017,pp.2.
17. Y. Koren, R. Bell, and C. Volinsky, “Matrix Factorization Techniques for Recommender Systems,” Computer (Long. Beach. Calif), Aug. 2009, vol. 42, no. 8, pp. 30–37.
18. Y. Koren, “Factorization meets the neighborhood,” in Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD 08, 2008, pp. 426.
19. D. D. Lee and H. S. Seung, “Learning the parts of objects by non-negative matrix factorization,” Nature, Oct. 1999, vol. 401, no. 6755, pp. 788–791.
20. R. Salakhutdinov and A. Mnih, “Probabilistic Matrix Factorization.,” in Advances in Neural Information Processing Systems (NIPS), 2008, pp. 1257–1264.
21. S. Seo, J. Huang, H. Yang, and Y. Liu, “Interpretable Convolutional Neural Networks with Dual Local and Global Attention for Review Rating Prediction,” in Proceedings of the Eleventh ACM Conference on Recommender Systems -RecSys ’17, 2017, pp. 297–305.
22. T. Donkers, B. Loepp, and J. Ziegler, “Sequential User-based Recurrent Neural Network Recommendations,” in Proceedings of the Eleventh ACM Conference on Recommender Systems - RecSys ’17, 2017, pp. 152–160.
23. C. Chen, M. Zhang, Y. Liu, and S. Ma, “Neural Attentional Rating Regression with Review-level Explanations,” in Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW ’18, 2018, pp. 1583–1592.
24. F. Costa, S. Ouyang, P. Dolog, and A. Lawlor, “Automatic Generation of Natural Language Explanations”, CoRR abs/1707.01561, Jul. 2017.
25. X. Chen, Y. Zhang, H. Xu, Y. Cao, Z. Qin, and H. Zha, “Visually Explainable Recommendation” , CoRR abs/1801.10288, Jan. 2018.
26. Manning, Christopher D., Mihai Surdeanu, John Bauer, Jenny Finkel, Steven J. Bethard, and David McClosky. 2014. “The Stanford CoreNLP Natural Language Processing Toolkit” In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations, pp. 55-60.