International Journal for Housing Science and Its Applications
Publish August 6, 2025. Volume 46, Issue 3 Pages 6919-6928

https://doi.org/10.70517/ijhsa463593

A Study on the Semantic Network Construction of Dress Metaphors in Ming Dynasty Novel Texts

Menghe Tian[1] and Xiangyang Bian[1,*]
[1] Collage of Fashion and Design, Donghua University, Shanghai, 200050, China
Corresponding authors: (e-mail: xiangyanbian@163.com).

Abstract Ming Dynasty novels are the golden period of Chinese literary development, in which dress metaphors, as an important means of literary expression, contain rich cultural connotation and social significance. This study focuses on dress metaphors in Ming Dynasty novel texts, and constructs a semantic network of dress metaphors through the improved TF-IDF algorithm, latent semantic analysis method, and semantic network analysis method to mine the hidden information in Ming Dynasty novel texts. The study constructed a corpus containing 4580 sentences, including 2677 non-metaphorical sentences and 1903 metaphorical sentences. The quantitative calculation of the dress words by LSA method reveals that the dress words with the highest correlation in the original and metaphorical meanings are the tonic and the wuzi cap, with the correlation of 0.114241 and 0.124387, respectively. The analysis of Water Margin shows that the red headscarf, as a symbol of the resistance against oppression of the Liangshan heroes, has the highest frequency of appearing throughout the book, with a frequency of 218 times; and the wuzi cap, which symbolizes the bureaucratic system, and tonic appear 188 and 182 times respectively, reflecting the twists and turns of the characters' fates in the work. The semantic network analysis further reveals the intrinsic connection between dress metaphors and the novel's themes, and shows how dress words carry social and cultural information and changes in characters' fates by means of metaphors. This study provides new computational methods and theoretical perspectives for analyzing the texts of Ming Dynasty novels, and helps to deepen the understanding of metaphorical expressions in Ming Dynasty literature.

Index Terms Ming Dynasty novels, dress metaphor, semantic network, latent semantic analysis, TF-IDF algorithm, text mining

I. Introduction

In the process of “humanization of nature” and “socialization of man”, clothing has been continuously developed and perfected, and its function is no longer limited to the practical significance of covering the body, keeping warm, and preventing various kinds of aggressions, but it has become a symbol of expression, and part of the discourse of the times [1]-[3]. As a symbol, dress is an outward manifestation of the wearer's social status, national identity, gender role, interests and desires as well as aesthetic habits [4]. And the depiction of dress in literary works has become the key to the writers' characterization and a symbol of the character's personality and destiny [5]. In the novel texts of the Ming Dynasty, we can see a lot of detailed descriptions of dress, which is undoubtedly a metaphorical dress discourse carefully constructed by the writers.

In the novel texts, the dress realizes both the covering of the body and the display of the beauty of the body, especially for women, the dress is an excellent tool for protecting personal privacy and displaying personal charm [6], [7]. In addition to clothes, some other decorations, such as gloves and hats, are also part of expressing a character's unique personality and self-worth [8], [9]. In addition, dress is also a symbol of social identity and status, and people of different classes and statuses will have different choices of dress [10]. Clothing becomes an important criterion for distinguishing people's identity, or people's identity status and personal value are shown only through materialized clothing, and individualized people are solidified into groups with different identities, and it is difficult to change this identity [11]-[13]. Due to the imaginative nature of novel creation, it is difficult to understand the full content of the dress metaphor by personal experience alone, and the information that exists in a large number of secular novels has significant research value. Therefore, digital research methods should be actively introduced to further excavate the costume metaphors widely present in novel texts [14], [15].

The Ming Dynasty was the golden period of the development of Chinese novels, which produced a large number of excellent novels, vividly presenting the rich and colorful humanistic landscape and social culture at that time. As the basic means of human survival, clothing in Ming novels not only has practical functions, but also carries rich social and cultural connotations and symbolic meanings. As an important means of literary expression, clothing

metaphor conveys specific cultural information and value judgment by linking the specific features of clothing with abstract concepts. In the novels of the Ming Dynasty, costume metaphors existed widely and in various forms, such as the “red scarf” in Water Margin, which symbolizes the spirit of resistance, and the “black hat”, which represents the power of the officialdom, which give the text a richer connotation and a more interpretative space through the way of metaphor. However, traditional literary research mainly relies on qualitative analysis, and there is a lack of systematic excavation and quantitative research on costume metaphors, which leads to certain limitations in grasping and understanding costume metaphors in Ming Dynasty novels as a whole. As a semantic way of generating words, clothing metaphors not only reflect the material culture of Ming social life, but also imply a specific social order and class structure. An in-depth analysis of dress metaphors in Ming novel texts can help reveal the hidden information in the texts and understand the unique expression and aesthetic characteristics of Ming literary creation. Currently, the development of computational linguistics and text mining technology provides new methods and perspectives for literary research. Quantitative analysis and semantic network construction of dress metaphors in texts through computational methods can more systematically and comprehensively grasp the distribution characteristics and expression functions of dress metaphors in Ming novels. Therefore, applying modern computational linguistics methods to the study of dress metaphors in Ming Dynasty novels can not only enrich the methodological system of traditional literary research, but also reveal the cultural connotation and artistic value of Ming Dynasty novels from a new perspective.

In this study, the improved TF-IDF algorithm, the latent semantic analysis method and the semantic network analysis method will be used to systematically study the dress metaphors in Ming Dynasty novel texts. Firstly, we obtain the text corpus of Ming Dynasty novels through Python crawler technology, and carry out preprocessing and keyword extraction; secondly, we quantify the dress metaphors by using the latent semantic analysis method, and analyze the semantic relevance of the dress words in the original and metaphorical sentences; then, we take Water Margin as an example, and apply the semantic network analysis method to construct the semantic network of the dress metaphors, and show the correlation between the dress metaphors by visualization; finally, we analyze the dress metaphors by the improved TF-IDF algorithm and semantic network analysis method. Finally, through the interpretation of the semantic network, the cultural meaning and social significance behind the dress metaphors are explored, and the hidden information in the Ming Dynasty novel texts is revealed. Through the comprehensive use of multiple computational methods, this study strives to analyze the expressive features and cultural connotations of the dress metaphors in Ming novels from the perspective of combining quantitative and qualitative methods, so as to provide new ideas and methods for the study of Ming literature.

II. Study design

The Ming Dynasty was the golden period of Chinese novels, and many excellent novels vividly presented the rich and colorful human culture at that time. Clothing is the most basic survival material of human beings, and metaphor is one of the more common semantic ways of generating words, and there are a large number of metaphorical words in Ming Dynasty novels. In this paper, we focus our research attention on the clothing metaphors in Ming Dynasty novels to explore the hidden information in Ming Dynasty novels.

II. A.Data sources and processing

All the data analyzed in this paper are textual data. Among them, the data of Chinese Ming dynasty novels come from online Ming dynasty literature websites. By using Python's Scrapy framework, given different initial URLs of Ming Dynasty literature, the required data can be crawled quickly, and then pre-processed, including mechanical compression, removal of invalid content, construction of deactivated dictionaries, etc., to finally obtain effective information.

II. B.Research methodology

II. B. 1) Improved TF-IDF algorithm

TF-IDF is a more effective keyword extraction algorithm, which is based on the word frequency and the size of the inverse text frequency to comprehensively determine the degree of importance of a word, the size of the word frequency in the political expository articles tends to reflect the importance of the degree of importance, and thus compared with TextRank, LDA theme algorithms, it is more reflective of its superiority [16]. The algorithm is divided into two parts: one is the TF algorithm; the other is the IDF algorithm, and in practical use, the TF is calculated as follows:

$$tf_i = \frac{n_{ij}}{\sum_k n_{kj}} \tag{1}$$

where n_{ij} denotes the frequency of occurrence of word i in document j, and after normalizing the word frequency, the denominator is the sum of the number of occurrences of each word in the statistical document.

The IDF is calculated as:

$$idf_i = \log\left(\frac{|D|}{1+|D_i|}\right) \tag{2}$$

where $|D|$ is the total number of documents in the document set, $|D_i|$ is the number of documents in the document set in which the word i occurs, and Laplace smoothing is used for the denominator plus one, to avoid the situation that there are part of the new segments that do not appear in the corpus and lead to the denominator being zero, and to enhance the robustness of the algorithm.

Then for the TF-IDF value of word i, there are:

$$tfidf_i = tf_i \times idf_i \tag{3}$$

However, the traditional TF-IDF ignores the problems caused by interclass concentration and intraclass dispersion, so it needs to be improved by introducing a correction factor η. The η is denoted as:

$$\eta = \frac{m_i}{m_i + k_i} \times \frac{m_i}{m_i + q_i} \tag{4}$$

where m and k are denoted as the number of contained words i in a certain document set and other documents, respectively, and q is the number of non-contained words i, and thus the correction coefficient can be expressed as the product of concentration and dispersion. The final improved TF-IDF formula is as follows:

$$tfidf_i' = tf_i \times idf_i \times \eta \tag{5}$$

II. B. 2) Latent Semantic Analysis Methods

In the potential semantic space constructed by using W-D matrix, each document can be regarded as a point in the semantic space with the dimension of keywords, and the distribution of documents and keywords in this space is definitely not random, but obeys a certain semantic structure. Similarly, these keywords can also be regarded as the corresponding points in the semantic space with the dimension of documents, and their distribution also obeys a certain semantic structure.

The implementation of the latent semantic analysis method mainly consists of three steps, i.e., constructing the $W-D$ matrix, singular value decomposition, and reconstructing the latent semantic space [17]. When constructing the latent semantic space model, factors such as the complexity of data computation, occupied storage space, consumed memory space, semantic expression ability and performance evaluation index of the model, and the ability to handle data sparsity will affect the effect of semantic search.

The traditional $W-D$ matrix uses the number of times a word occurs in a document to represent the elements of the matrix, and in general a word only occurs in a small portion of the document, while the number of times it occurs in most of the documents is 0. Therefore, the $W-D$ matrix generated based on the frequency of the word contains a large number of 0-valued elements, which inevitably creates the problem of sparsity of the matrix, which has an impact on the processing of large-volume datasets that is fatal. Moreover, different words in a document contribute differently to the semantic information content of the document, and two words with the same number of occurrences in the same document do not necessarily have the same role in distinguishing the document, so it is not reasonable to use word frequency to measure the semantic information content of words.

Later, TF-IDF weighting method was used to define the elements of the $W-D$ matrix, and the TF-IDF method usually filters out common words and retains important words. The premise of this method is that if a word occurs more frequently in a given document and less frequently in others, then it has a greater ability to distinguish documents, and if it occurs in fewer documents, then it has a greater ability to distinguish documents. The essence of this method is to suppress the weighting of noise, but at the same time, it suppresses the weight adjustment effect of those high-frequency feature words that represent the important features of the document, so in some cases, the accuracy of this method is not high.

To summarize, in latent semantic analysis, singular value decomposition of the original $W-D$ matrix generated according to word frequency or TF-IDF method will lose a large amount of semantic information, which may lead to bias in the accuracy of the final semantic search.

Since the potential semantic analysis method needs to achieve an effect of information approximation and semantic extraction, which coincides with the idea of weights calculation, the following weights calculation method has been proposed:

$$W(i,j)=W_L(i,j)\times W_G(i) \tag{6}$$

where $W(i,j)$ denotes the weight of word i in document j, $W_L(i,j)$ is the local weight of word i in document j, and $W_G(i)$ denotes the global weight of word i in the whole document set.

This weight calculation method only considers the contribution of the information content of words in reflecting the semantic relationship of documents, while ignoring the contribution of the relationship between documents and the semantic information content of documents.

In order to reflect the differences in semantic contributions of different words and documents to construct the potential semantic space and the role of information generalization, this paper adopts the following new weight calculation method:

$$W(i,j)=W_L(i,j)\times W_G(i)\times D_G(j) \tag{7}$$

The Λ represents the original $W-D$ matrix. The new weight calculation method adds the factor $D_G(j)$, the document global weight, to the original method, which indicates the role played by the document in distinguishing the amount of semantic information provided by the words. The purpose of this is to consider the role of the document in distinguishing the semantic information as a trade-off factor that can extract the semantic information of the document to a greater extent.

1) Word local weights:

$$W_L(i,j)=\log_2\left(wf_{ij}+1\right) \tag{8}$$

wf_{ij} denotes the number of times a word i appears in a document j, and adding it to the constant 1 is to prevent $W_L(i,j)$ from being meaningless when wf_{ij} is zero. The logarithmic base here is generally taken as 2. In practice, the value of the logarithmic base can be decided according to the sparsity and distribution of non-zero elements of the $W-D$ matrix and the average length of the document.

This method of taking the logarithm of word frequency to define the local weights of words can avoid words with too high word frequency masking the contribution of words with too low word frequency to the potential semantic space.

2) Word global weights:

$$W_G(i)=\frac{H(doc)-H(doc\mid word_i)}{H(doc)}=1-\frac{H(doc\mid word_i)}{H(doc)} \tag{9}$$

here according to the definition of conditional entropy:

$$H(doc\mid word_i)=-\sum_j p(j\mid i)\times\log_2 p(j\mid i) \tag{10}$$

And $H(doc)$ is a constant, by the nature of entropy it can be proved that $H(doc)\le\log_2 n$, according to many experiments found that the value of $H(doc)$ is approximately equal to the value of $\log_2 n$. Therefore Eq. (10) can be rewritten as:

$$W_G(i)=1+\frac{\sum_j p(j\mid i)\times\log_2 p(j\mid i)}{\log_2 n} \tag{11}$$

where $p(j\mid i)=\frac{wf_{ij}}{gf_i}$, denotes the probability of "document j appears if word i appears"; gf_i is the number of times word i appears in the whole document set.

This definition is derived from the idea of entropy weight, the global weight of a word measures the size of the semantic contribution of a word i to the entire document set, which to some extent represents the size of the role of words in distinguishing documents.

3) Document global weights:

$$D_G(j)=\frac{H(word)-H(word\mid doc_j)}{H(word)}=1-\frac{H(word\mid doc_j)}{H(word)} \tag{12}$$

Here:

$$H\left(word \mid doc_j\right) = -\sum_i p(i \mid j) \times \log_2 p(i \mid j) \tag{13}$$

where $p(i \mid j) = \frac{wf_{ij}}{dw_j}$, denotes the probability of "the occurrence of word i if document j occurs"; and dw_j is the total number of words contained in document j, i.e., the document length. And:

$$H(word) = -\sum_{i=1}^{m} \frac{gf_i}{gf} \times \log_2 \frac{gf_i}{gf} \tag{14}$$

where $gf = \sum_{i=1}^{m} gf_i$, the total number of occurrences of all words in the document set. The semantics of a document consists of the amount of semantic information of all the words it contains, and the semantics of a word is closely related to the topic of the document it is in

II. B. 3) Semantic Network Analysis

Semantic network analysis is used to analyze the dress metaphor in Ming Dynasty novels [18]. The whole is divided into two steps: the first step is still the extraction of keywords; the second step is the co-occurrence analysis, according to the importance of the keywords to get the co-occurrence matrix, and then according to the co-occurrence matrix to further understand the relationship between the words and the degree of affinity.

II. C.Research tools

The tools used in this study are Python, ROST CM 6 and Ge-phi 0.9.2.The data crawling and preprocessing is done in Python and keyword extraction is done using the improved TF-IDF algorithm, followed by semantic network analysis using ROSTCM 6, and finally, the co-homogeneous matrices obtained from the semantic network analysis are visualized using Gephi 0.9.2.

III. Quantitative Calculation of Dress Metaphors in Ming Dynasty Fiction Texts

In this chapter, we will apply the Latent Semantic Analysis method (hereinafter referred to as the "LSA method") to quantify the dress metaphors in Ming Dynasty novel texts.

III. A. Clothing Metaphor Corpus Construction and Preprocessing

Due to the late start of metaphor computing research, metaphor corpus resources, especially computer-oriented metaphor corpus, are still very scarce. Metaphor research is based on real metaphor corpus, we construct our own metaphor corpus of Ming Dynasty novels according to the definition of metaphor category.

After obtaining the original text corpus, there are many dirty data mixed in the corpus, such as missing text and text repetition. By removing these obvious dirty data, not only can the data processing speed be accelerated, but also can improve the accuracy of text analysis, so it is necessary to pre-process the data. The invalid spaces, line breaks and other illegal characters such as network symbols are removed from the text by regular expressions; then the duplicated text in the corpus is eliminated because the duplicated text noise will be inaccurate in the process of computation, which affects the normal analysis results; finally, the preprocessed text corpus is manually labeled with metaphorical and non-metaphorical expressions by expert readers.

After the above preprocessing process, the preprocessed original corpus is obtained, as shown in Table 1. The number of sentences obtained is 4580, of which 2677 and 1903 are non-metaphorical and metaphorical sentences respectively.

Table 1: The pre-processed original corpus

Type	Quantity
Number of sentences	4580
Non-metaphorical sentences	2677
Metaphorical sentences	1903

III. B. Real Word Document Matrix Generation

Using Python and the CountVectorizer class in the scikit-learn package, we construct a computer-processable data structure for the extracted real words, and use the fit_transform function to count the word frequencies and generate

word frequency document data. The resulting word frequency document data is a very sparse matrix, where each row represents a document, each column represents a real word after preprocessing, and the value of the cross term of the rows and columns represents the frequency of the real word appearing in the document. The size of the word frequency document matrix after the processing of the vector space model for the vernacular sentence and metaphorical sentence is shown in Table 2. As can be seen from the table, the size of the real word document matrix is 1325*808 in the sense and 1022*711 in the metaphor after the processing of the vector space model.

Table 2: Content word document data structure

Type	Number of documents	Number of substantive words
The original meaning	808	1325
Metaphor	711	1022

III. C. Calculation of potential semantic analysis of real word relevance

The high-dimensional and sparse word frequency document matrix generated by the vector space model is downsized by SVD, and the dimension K is set to 10 in this paper.According to the results of real word relevance calculation by LSA, the real word relevance list of keywords is extracted, and the real word relevance list of LSA is shown in Table 3. It can be seen that at this time the highest relevance of the dress words in the original meaning and metaphorical meaning is the tonic, the ossuary, the relevance is 0.114241, 0.124387, respectively.

Table 3: List of LSA notional word correlations

Number	The original meaning		Metaphorical meaning	
	The real word	Degree of correlation	The real word	Degree of correlation
1	Buzi	0.114241	Black gauze cap	0.124387
2	Black gauze cap	0.093618	Buzi	0.094732
3	Horseface skirt	0.089039	Red embroidered shoes	0.076872
4	Straightening	0.06627	Jade belt	0.050256
5	Red embroidered shoes	0.060774	Straightening	0.04512
6	Jade belt	0.055629	Square towel	0.040829
7	Square towel	0.050971	Horse Skirt	0.037789
8	Bijia	0.049041	Broken clothes	0.037504
9	Cassock	0.047808	Bijia	0.034973
10	Foot-binding cloth	0.04217	Foot-binding cloth	0.032467

III. D. Real word correlation Person calculation

Join irrelevant 50 noise sentences together as a dataset for Pearson real word relevance calculation and extract the keyword real word relevance list. The keyword real word relevance list is shown in Table 4. In the native and metaphorical Pearson keyword real word relevance lists after adding noise, the most relevant dress words are still mending and ursine, with the relevance of 0.091884 and 0.102054, respectively.

Table 4: Pearson content word relevance list

Number	The original meaning		Metaphorical meaning	
	The real word	Degree of correlation	The real word	Degree of correlation
1	Buzi	0.091884	Black gauze cap	0.102054
2	Black gauze cap	0.081901	Red embroidered shoes	0.078659
3	Straightening	0.067401	Buzi	0.068883
4	Red embroidered shoes	0.063531	Jade belt	0.062295
5	Jade belt	0.061801	Straightening	0.058054
6	Horse Skirt	0.061394	Square towel	0.053318
7	Square towel	0.05833	Broken clothes	0.053288
8	Cassock	0.05756	Bijia	0.050868
9	Bijia	0.057534	Foot-binding cloth	0.049669
10	Cassock	0.053025	Bijia	0.048397

III. E. Comparison of Latent Semantic Analysis and Pearson Real Word Correlation

The list of keyword real word relevance extracted by LSA calculation results and the list of keyword real word relevance extracted by Pearson calculation results appear very similar to the real word results. The two methods with different relevance of real words with the same meaning are visualized and analyzed, as shown in Fig. 1.The LSA and Pearson calculation of real word relevance are both decreasing in the overall trend, but the distinction between the data results obtained by the LSA calculation method is more obvious. Comparison of these two sets of data indicates that LSA varies significantly in correlation, with high correlation values for close relationships and low correlation values for distant relationships. It is further shown that the LSA calculation method is more obvious than the results of Pearson method, which indicates that LSA has better results of real word similarity calculation.

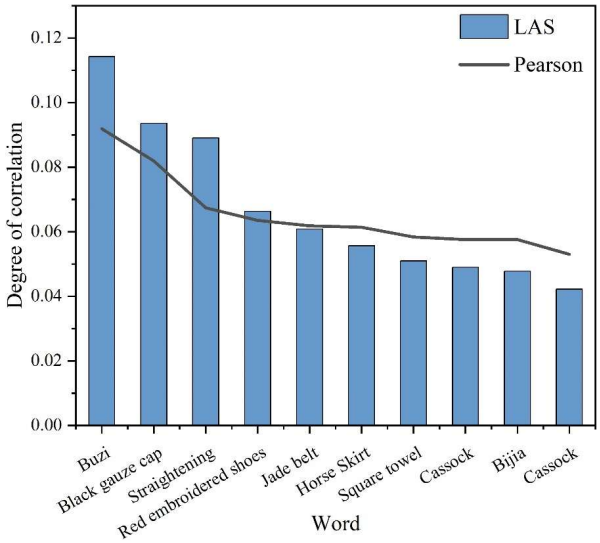

Figure 1: Comparison of LSA and Pearson calculation

III. F. Statistics of real words occurring simultaneously in onomatopoeic and anaphoric sentences

The LSA calculation method was applied to the vernacular and metaphorical sentences, and the processing results revealed that the same real words appeared in the vernacular and metaphorical sentences at the same time, and the real word relevance of these same real words in the vernacular and metaphorical sentences was different. In order to clearly show the effect of real words appearing simultaneously in the vernacular and metaphorical sentences on the vernacular and metaphorical, the results of LSA real word relevance calculation were used for data analysis. Every 100 real words as a group and take the first 10 groups of data, respectively, denoted by A1~A10, to count the number of the same real words appearing in the vernacular and metaphor, as shown in Figure 2. From the figure, it can be seen that there are 405 real words in the top 1000 real words with the highest degree of relevance in the apparel vocabulary appearing in the metaphor corpus. Among them, 85 real words in the top 100 real words with the highest relevance (the first group) appear in the metaphor corpus, and the remaining 15 real words may be the feature real words that distinguish the original meaning from the metaphor.

As the correlation decreases, the number of the same real words in the vernacular and metaphorical corpus also decreases, indicating that the real words with high correlation are the high-frequency words constituting the documents, however, these real words do not play an obvious role in distinguishing between vernacular and metaphorical. In addition, as the correlation decreases, the number of common real words in the vernacular and metaphorical corpus gradually decreases, and the number of differentiated real words gradually increases. It shows that differentiated real words are important clues for distinguishing between vernacular and metaphorical sentences, and although the correlation decreases, the differentiation increases, a finding that provides new ideas and directions for metaphorical research.

IV. Textual Analysis of Dress Metaphors in Ming Dynasty Fiction Texts

In this chapter, we will take the Chinese Ming Dynasty novel Water Margin as an example, and use the improved TF-IDF algorithm and semantic network to analyze the metaphorical textual information of Ming Dynasty novels' dresses, and mine the hidden information in the text of Ming Dynasty novels.

IV. A. Statistics on basic text information

By writing code, you can extract the information you need, such as entering relevant code, and count the frequency of "black yarn hat" and "complement" in the whole text, as shown in Figure 3. The black gauze hat is an official hat made of official uniforms, and the complement is a square embroidered pattern on the front chest of the official uniform. It can be seen that in the last 40 episodes, Song Jiang led the heroes of Liangshan to surrender to the imperial court, and the number of "black yarn hats" and "complements" symbolizing the power order and the hierarchy of the bureaucratic system increased significantly, and the plot of the whole work took a turn, and the fate of the heroes of Liangshan gradually moved towards tragedy.

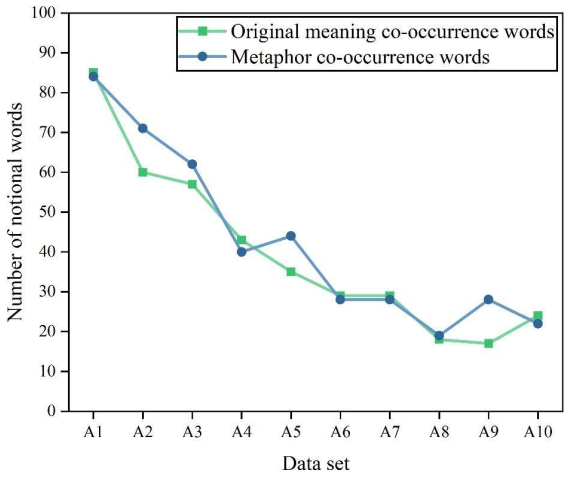

Figure 2: The distribution of original meaning and metaphorical co-reality words

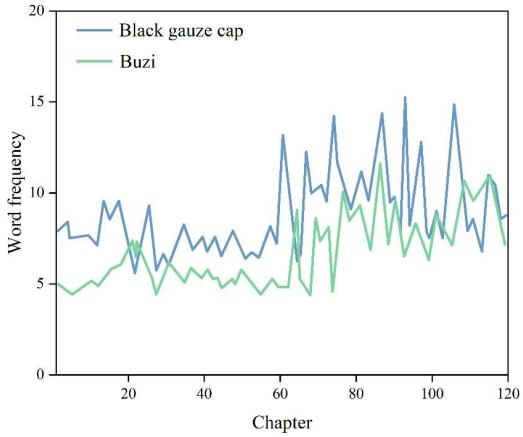

Figure 3: The frequency of 'Black gauze cap' and 'Buzi'Buzi

IV. B. Text word frequency statistics

Count the number of clothing metaphor words that appear in the whole book of "Water Margin", as shown in Table 5. It can be seen that the word red turban appears the most frequently, reaching 218 times, and is a symbol of the establishment of a rebellious community by the heroes of Yangshan through uniform clothing. As symbols of the imperial bureaucratic system, the "black gauze hat" and "complement" appeared 188 times and 182 times. The brocade jacket and bast shoes appear 195 times and 112 times respectively, and the comparison between the two is to reveal the comparative performance of the sudden change of the heroes of Liangshan from the bottom to the uprising and the sudden wealth. As a symbol of punishment, the gold seal appears 125 times, while the embroidery, as a symbol of the identity of the heroes of Liangshan who used their bodies to transform their own rivers and lakes, appears 192 times. The costume metaphors in Water Margin further deepen the thematic expression of the book.

Table 5: Word frequency statistics

Serial number	Word	Frequency
1	Red scarf	218
2	Jinjia	195
3	Flower embroidery	192
4	Black gauze cap	188
5	Buzi	182
6	Gold seal	125
7	Straightening	122
8	Leopard skin	118
9	Hosta flower	115
10	Hemp shoes	112

IV. C. Semantic network analysis of clothing metaphors

In this section, the lexicon of Python will be used for word segmentation and stored as a csv format file, and then the file will be imported into ROST CM 6 for semantic network analysis, and then the co-occurrence matrix obtained from semantic network analysis will be imported into Gephi 0.9.2 for visualization, and the results are shown in Fig. 4. The node size and the thickness of the edges of a semantic network correspond to the importance and co-occurrence of the words in that network, respectively, while the magnitude of the centrality of a word can be measured by the number of words connected to it. As can be seen from the above chart, the “red scarf” has the highest degree of centrality, indicating that it is a symbol of Liang Shan heroes' resistance to oppression, which is present in almost the entire book. Among the words connected with "red turban", the importance and centrality of "brocade jacket", "gold seal", "direct fix" and "flower embroidery" are close behind. Through the metaphor of clothing, we can see that in the whole book of Water Margin, the Liang Shan heroes rebel against oppression and pursue loyalty and righteousness, and the relationship between the Liang Shan heroes and the court fluctuates and fluctuates along with the development of the plot, with loyalty and betrayal, heroism and personal destiny intertwined with each other. The Liang Shan heroes wear red bandanas as their identity because they are gathered in Liang Shan for righteousness, while in the later period they return to the imperial court for loyalty and wear official robes with straight embellishments.

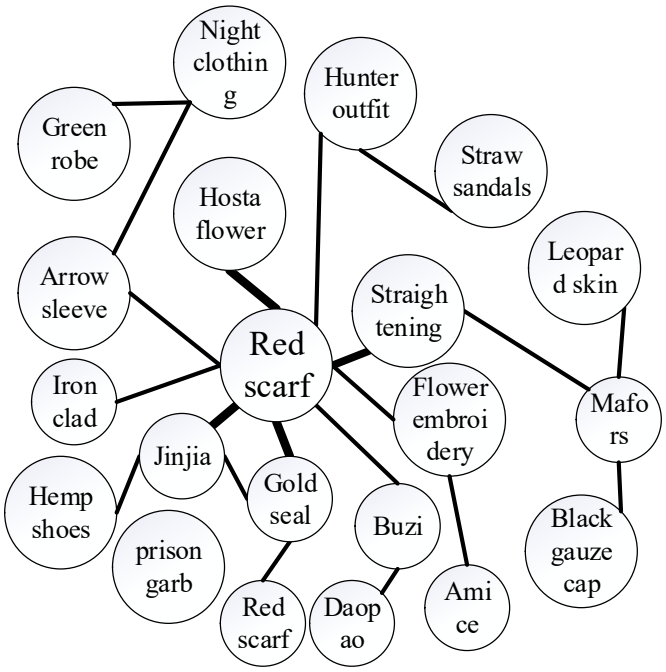

Figure 4: Semantic network of clothing metaphor

V. Conclusion

The study reveals the expressive features and cultural connotations of dress metaphors in Ming Dynasty novels by studying the semantic network construction of dress metaphors in Ming Dynasty novel texts. The results of the latent semantic analysis show that in the Ming novel texts, the most relevant dress words in the original meaning and metaphorical meaning are mending (0.114241) and wu sha hat (0.124387), respectively, and these high-frequency dress words, as the symbols of identity status and order of power, profoundly reflect the hierarchical system and value orientation of the Ming society. An analysis of Water Margin reveals that the “red turban”, which is the symbol of resistance of the Liangshan heroes, appears most frequently (218 times), while the “black hat” and “mending son”, which symbolize the bureaucracy, appear 188 and 182 times respectively, which reflects the twists and turns of the characters' destinies and the tragic colors in the novel. The semantic network analysis of clothing metaphors further reveals that “red turban” has the highest degree of centrality in the network, and is closely associated with the words “brocade jacket”, “golden seal”, “straight embellishment” and “flower embroidery”, which together build up the thematic expressions of loyalty and righteousness, resistance and submission in the novel. The semantic network analysis method provides a new perspective for mining the cultural connotation behind the dress metaphor, and compared with the traditional text analysis method, it can more systematically and comprehensively grasp the distribution characteristics and functions of the dress metaphor in Ming Dynasty novels. The dress metaphors in Ming novels not only reflect the material and cultural phenomena, but also carry specific social order and cultural values. Through the construction of the semantic network of dress metaphors, we can deepen our understanding of the artistic characteristics and cultural connotations of Ming novels, and provide new methodological support and theoretical perspectives for the study of ancient Chinese literature.

References

[1] Song, C., Zhao, H., Men, A., & Liang, X. (2023). Design Expression of" Chinese-style" Costumes in the Context of Globalization. Fibres & Textiles in Eastern Europe, 31(2).

[2] Hester, N., & Hehman, E. (2023). Dress is a fundamental component of person perception. Personality and Social Psychology Review, 27(4), 414-433.

[3] Alcantara, A. C., & Michalack, D. L. (2022, June). The metaverse narrative in the matrix resurrections: A semiotic analysis through costumes. In Meeting of research in music, arts and design (pp. 260-268). Cham: Springer International Publishing.

[4] Figueiredo, C. M., & Cabral, A. (2018). Costume and characters’ construction in cinematographic fiction and drama. In Textiles, Identity and Innovation: Design the Future (pp. 73-78). CRC Press.

[5] Lerer, S. (2020). Costumes for our present: How every age rewrites the old into the new. TLS. Times Literary Supplement, (6116), 26-28.

[6] Wei, L. (2018). A preliminary study of Mongol costumes in the Ming dynasty. Social Sciences in China, 39(1), 165-185.

[7] Lu, C., & Ismail, N. H. (2024). The Characteristics of the Design of Ming Dynasty Palace Festival Costumes. Asian Journal of Research in Education and Social Sciences, 6(1), 794-806.

[8] Li, D., Liang, H. E., & Zhang, A. (2022). Flaunting and Consumption: Women’s Headdress in the Ming Dynasty of China. Fashion Theory, 26(3), 383-398.

[9] Zhu, Y. (2021). Rank badges of official costumes of Ming and Qing dynasties from the perspective of social semiotics. Language and Semiotic Studies, 7(2), 121-143.

[10] Tian, M., & Bian, X. (2025). Computational Linguistic Interpretation and Modeling of Dress Metaphors in Ming Dynasty Novel Texts. J. COMBIN. MATH. COMBIN. COMPUT, 127, 6909-6925.

[11] Tajuddin, F. N. (2018). Cultural and social identity in clothing matters “different cultures, different meanings”. European Journal of Behavioral Sciences, 1(4), 21-25.

[12] Lu, C., & Ismail, N. H. (2023). THE DESIGN CHARACTERISTIC OF LATE 15TH CENTURY MING DYNASTY COSTUMES IN THE JIANGNAN REGION. Jurnal Tuah, 4(1), 179-188.

[13] Li, X., Liu, K., & Gao, F. (2023). Research on the restoration of costumes in the paintings of ladies in the Ming Dynasty. In Machine Learning, Multi Agent and Cyber Physical Systems: Proceedings of the 15th International FLINS Conference (FLINS 2022) (pp. 656-663).

[14] Chuwei, Z., & Yunmei, S. H. I. (2023). Metaphor Recognition Method based on Graph Neural Network. International Journal of Advanced Computer Science and Applications, 14(1).

[15] Zhao, N., & Zhao, W. (2024). Enhancing Metaphor Recognition of Literary Works in Applied Artificial Intelligence: A Multi-Level Approach with Bi-LSTM and CNN Fusion. Applied Artificial Intelligence, 38(1), 2413817.

[16] Lu Jiaxin. (2025). Text vectorization in sentiment analysis: A comparative study of TF-IDF and Word2Vec from Amazon Fine Food Reviews. ITM Web of Conferences,70.

[17] B. Sindhu,A. Bhaskar,G. Yugesh,S. Reshma & B. Rohit. (2025). Enhancing Educational Video Discovery Using Advanced Latent Semantic Analysis. Procedia Computer Science,252,784-795.

[18] Tang Wujun,Chung Jiwon & Koo Sumin. (2025). Key terms and topics of muscle-supportive and posture-corrective wearable robots for older adults using text mining and semantic network analyses. Research Journal of Textile and Apparel,29(2),389-407.