International Journal for Housing Science and Its Applications
Publish August 10, 2025. Volume 46, Issue 4 Pages 1745-1754

https://doi.org/10.70517/ijhsa464152

An Analysis of the Impact of Chinese Feminism on Changes in the Social Status of Women in Modern and Contemporary China Based on the Materialist Historical Perspective

Yihan Liu[1,*]
[1] School of Marxism, Shenyang City University, Shenyang, Liaoning, 110112, China
Corresponding authors: (e-mail: 18540234689@126.com).

Abstract Based on the theoretical framework of historical materialism and with the help of text mining technology, this study systematically examines the influence mechanism of Chinese feminism on the change of women's social status in modern times. Three hundred texts on modern women's issues between 1919 and 1949 are selected for word frequency statistics, which are combined with the LDA topic model to reveal the relevance of key themes. The experimental and control groups are divided to verify the influence of feminism on marital autonomy and family decision-making mode. Analyze the specific factors affecting women's social status advancement through multiple regression analysis. Years of education, occupational participation, legal awareness and annual family income have significant positive effects on women's social status satisfaction. Among them, the influence of years of education is the most prominent (β=0.32), the standardized coefficient of occupational participation (β=0.28) is the second most important, and the significant influence of legal cognition score (β=0.25) indicates that the awakening of the awareness of the rights is an important psychological mechanism for the improvement of women's social status. Annual household income has a positive but weak effect (β=0.18), while age fails the significance test. Overall, the model explains about 48% of the variance in social status satisfaction, providing a new analytical perspective for understanding the historical logic of modern female emancipation.

Index Terms materialism, feminism, text mining techniques, LDA topic model, multiple regression analysis

I. Introduction

The issue of women's emancipation is a universal world problem and has always been a great concern for feminists or academics [1], [2]. Marxist feminism, which combines Marxism and theories related to women's liberation, is one of the main schools of Western Marxism [3]. It draws on and cites Marxist theories about economy, class and other aspects, analyzes the fundamental causes of contemporary women's oppression from the perspective of the Marxist materialist view of history, and proposes an important way for women's liberation [4]-[6]. Marxist feminists advocate the use of class analysis to reveal the root causes of women's oppression, to recognize women's social status and social role, and to explore the path to women's emancipation [7]-[9]. It is a fundamental point of view of Marxist feminists that the particular mode of production and patriarchy under the capitalist system is the root cause of women's oppression and exploitation [10].

What can be glimpsed is that since the emergence of the issue of women's emancipation till date, the status of women in the society has changed drastically [11]. Although the status of women in society has been greatly improved, and even to some extent, men and women have basically realized equality between men and women, the issue of women's emancipation still has a long way to go [12]-[14]. Using Marxist feminism as a research tool to analyze the changes in the social status of women in modern China and to scientifically guide the difficulties encountered in the process of Chinese women's cause is conducive to promoting the comprehensive and healthy development of China's women's cause [15]-[18].

In this paper, we take the modern women's issue texts from 1919-1949 as the research object, and quantify the modern women's issue texts based on statistical Chinese word segmentation and LDA topic model. Extract text high-frequency words and identify core themes. Analyze the high-weighted themes to reveal the evolution logic of feminist discourse. Set up a control group to examine the influence of feminist concepts on Chinese women's marital autonomy and family decision-making patterns. Utilizing multiple regression modeling to explore the role of years of education, occupational participation, and legal knowledge on women's social status satisfaction.

II. Text mining method based on LDA topic modeling

II. A. Statistical-based Chinese Segmentation Methods

In the era of frequent web interactions, new words and open-ended intents keep emerging, and it is difficult to cover all the vocabulary only by relying on dictionaries, and statistically based segmentation methods have become mainstream. The core idea of the method is to consider each word as a combination of the smallest unit characters, and if a consecutive string of Chinese characters appears more often and more frequently in different texts, the higher the probability that the string is a word. Therefore, the reliability of words can be measured by using the frequency of adjacent occurrences of words and characters, by counting the frequency of occurrence of adjacent Chinese character combinations in the corpus, and when the frequency exceeds a certain threshold, it is considered that the string combination can be classified as a word.

(1) N-gram model

N-gram model is a language model, also known as N-gram model, widely used in tasks such as lexical annotation, sentence segmentation, and text classification. It is a probability-based discriminative model for predicting the probability of occurrence of a sequence of words in a text. Thus, given a sequence of words, the model output is the joint probability of all the words in this sequence. A probabilistic interpretation of the principle is that for a string w of length n, assuming that w consists of a sequence of words $w_1 w_2 \dots w_n$, the probability value of its occurrence can be computed using the chain rule:

$$\begin{aligned} P(w) &= P\left(w_1, w_2, \cdots w_n\right) \\ &= P\left(w_1\right) P\left(w_2 \mid w_1\right) P\left(w_3 \mid w_1, w_2\right) \cdots P\left(w_n \mid w_1, w_2, \cdots, w_{n-1}\right) \end{aligned} \tag{1}$$

Observing equation (1), it can be seen that when the text is too long, the calculation of each term of the formula from the third term onwards is very difficult. The N-gram algorithm assumes that the occurrence of each word is only related to the $N-1$ words before it, which greatly reduces the difficulty of the calculation. Neglecting the effect of words with distances greater than or equal to N, the probability value of the occurrence of the string w can be simplified as:

$$P\left(w_i \mid w_1, w_2, \cdots, w_{i-1}\right) \approx P\left(w_i \mid w_{i-(N-1)}, \cdots, w_{i-1}\right) \tag{2}$$

When $N=1$, it is called a unitary model, when the probability representation of the whole sentence is:

$$P\left(w_1, w_2, \cdots w_n\right) = P\left(w_1\right) P\left(w_2\right) \cdots P\left(w_n\right) \tag{3}$$

When $N=2$, known as the binary model, the current word depends only on the word that precedes it, at which point the probabilistic representation of the entire sentence is:

$$P\left(w_1, w_2, \cdots w_n\right) = P\left(w_1\right) P\left(w_2 \mid w_1\right) \cdots P\left(w_n \mid w_{n-1}\right) \tag{4}$$

Obviously, when N is larger, the model retains richer word order information, but the computational cost is also higher.

(2) Hidden Markov Model (HMM)

HMM is a probabilistic model based on time series, which is used to deal with the task of word segmentation when it is transformed into the problem of labeling sequences in strings. In HMM, there exists a Hidden Markov Chain that randomly generates a sequence of states, each of which in turn generates an observable output, forming a sequence of observations. Thus, the HMM model consists of two sequences: the state sequence (hidden) and the observation sequence (visible). In the case of the disambiguation task, the state sequence corresponds to the labeling of each character, while the observation sequence corresponds to the input text string. With the HMM, the most probable state sequence for a given observation sequence can be computed, thus realizing text disambiguation. The basic principle of the HMM model is shown in Fig. 1.

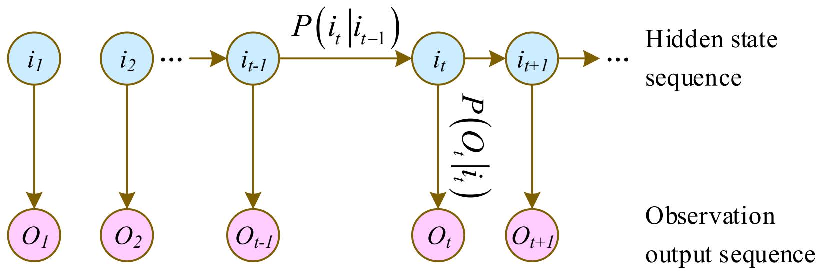


Figure 1: Fundamentals of the HMM model

where I denotes the state sequence of length T and O is the observation sequence corresponding to the state sequence I:

$$I=\{i_1,i_2,\dots,i_T\} \tag{5}$$

$$O=\{o_1,o_2,\dots,o_T\} \tag{6}$$

Hidden Markov models have two basic assumptions:

Markov assumption: it is assumed that the state of the Hidden Markov Chain at any moment t depends only on the previous moment $t-1$, independently of the states and observations at other moments, and independently of the moment t. Namely:

$$P(i_t \mid i_{t-1},o_{t-1},\cdots,i_1,o_1)=P(i_t \mid i_{t-1}),t=1,2,\cdots,T \tag{7}$$

Observation Independence Assumption: it is assumed that an observation at any moment depends only on the state of the Markov chain at that moment, independent of other observations and states, i.e.:

$$P(o_t \mid i_T,o_T,i_{T-1},o_{T-1},\cdots,i_{t+1},o_{t+1},i_t,o_t,\cdots i_1,o_1)=P(o_t \mid i_t) \tag{8}$$

The process of Chinese word splitting is to solve for the most likely sequence of hidden states given a sequence of observations, with $O=\{o_1,o_2,\dots,o_T\}$ denoting the input sentence and $I=\{i_1,i_2,\dots,i_T\}$ denoting the output label, the ideal output that is:

$$\max=\max P(i_1,i_2,\cdots,i_T \mid o_1,o_2,\cdots,o_T) \tag{9}$$

obtained using Bayesian formula:

$$P(i \mid o)=\frac{P(i,o)}{P(o)}=\frac{P(o \mid i)P(i)}{P(o)} \tag{10}$$

where O is the given input and $P(o)$ is a constant that can be ignored. Thus maximizing $P(i \mid o)$ is equivalent to maximizing $P(o \mid i)P(i)$, which can be obtained according to Eq. (7) and Eq. (8):

$$\begin{gathered}P(o \mid i)P(i)\sim P(o_1 \mid i_1)P(i_2 \mid i_1)P(o_2 \mid i_2) \\ P(i_3 \mid i_2)\cdots P(o_n \mid i_n)P(i_n \mid i_{n-1})\end{gathered} \tag{11}$$

The HMM model is determined by the initial state probability vector π, the state transfer probability matrix A, and the observation probability matrix B, which is expressed in mathematical notation as $\lambda(A,B,\pi)$.

Where, the initial state probability A refers to the probability that each state I, is distributed at the initial moment, the state transfer probability matrix B refers to the probability that the two states i_{t1}, i_{t2} are transitioning to each other $P(i_{t2} \mid i_{t1})$, and the observation probability matrix is The probability $P(o_t \mid i_t)$ of the occurrence of a different observation O_t in each state I_t. The Hidden Markov Chain is determined by the state transfer matrix A with the initial state probability vector π, and the generation of the observation sequence is determined by the combination of the observation probability matrix B and the state sequence O.

II. B.LDA Thematic Model Construction and Parameter Estimation

II. B. 1) Text vectorization

As the smallest unit of understanding natural language, the high-quality vector representation of words can help us solve various natural language processing tasks, so a large number of research works have proposed vectorization methods and models for words, and breakthroughs have been made. In this paper, we use Word2Vec to vectorize text, Word2Vec is an unsupervised shallow neural network language modeling algorithm, its advantage is not only to represent words into vectors, but also to learn the inter-semantic relations; it is able to learn to infer unseen words; and because of the use of shallow neural networks, the computation speed is fast. The main mechanism is that the meaning of any word can be represented by the words of its contextual environment, and the one-hot encoded vectors are used as the original input layer training samples, and the weight matrix is adjusted through continuous learning to realize its mapping to dense continuous vectors.

Word2Vec has two training models, Continuous Bag of Words (CBOW) model and Skip-Gram model. In this paper, we choose the Word2Vec tool provided in the Genism library, which belongs to the CBOW model, which is fast and efficient for training. The learning objective of the CBOW model training process is to maximize $P(Context(w) \mid w)$, the context window and word probability as shown in Fig. 2, the process of which is to preprocess a corpus of previous participles, and then convert the original training corpus into an iterator storing sentences, with the sentence iterators as the inputs passed into Gensim's Word2Vec's module for training, the steps are briefly described as follows:

(1) Policy text preprocessing: Jieba split word;

(2) Transform the original training corpus into a sentence iterator;

(3) Apply the Word2vec object in Gensim for training.

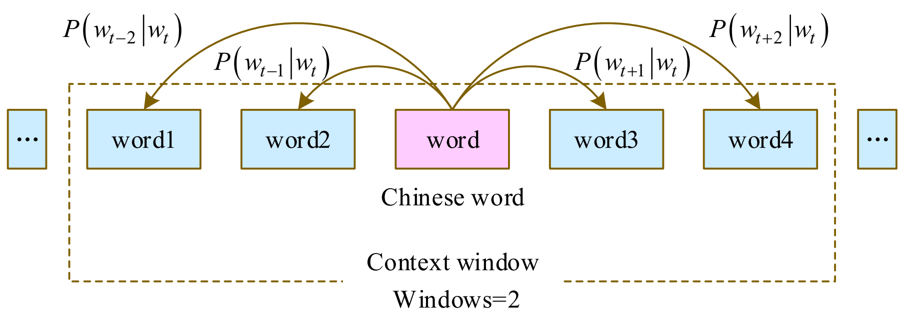


Figure 2: Probabilistic prediction

II. B. 2) Model construction and parameter estimation

The LDA topic model is essentially a probabilistic model, and the principle of the model is shown in Figure 3. The nodes in the figure represent random variables, solid nodes are observed variables, and null nodes are hidden variables; directed edges represent probabilistic dependencies; rectangles represent repetitions, and numbers in the panels represent the number of repetitions.

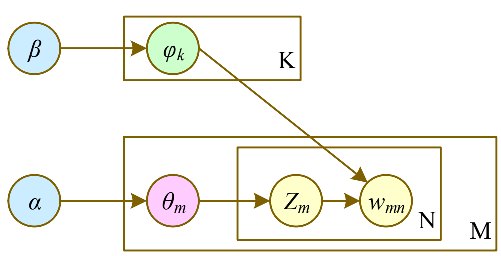


Figure 3: Principle of the LDA model

(1) Elements of the model

The LDA model uses three sets, the set of words $W=\{\omega_1,\cdots\omega_v,\cdots,\omega_V\}$ where ω_v is the v th word, $v=1,2,\cdots,V$, and V is the number of words. The set of texts $D=\{w_1,\cdots,w_m,\cdots,w_M\}$, where w_m is the first m th text, $m=1,2,\cdots,M$, M is the number of texts, $w_m=\left(\omega_{m1},\cdots,\omega_{mn}\cdots,\omega_{mN_n}\right)$, and the text w_m is a sequence of words,

ω_{mn} is the n th word of the text, $n=1,2,\cdots,N_m$, and N_m is the number of words in the text; the set of topics $Z=\{z_1,\cdots,z_k\cdots z_k\}$, where z_k is the k th topic, and $k=1,2,\cdots,K$, K is the number of topics.

Each topic is determined by a one-word conditional probability distribution $p(\omega|z_k)$, $p(\omega|z_k)$ which obeys a polynomial distribution with parameters φ_k, φ_k parameterized by a Dirichlet distribution with hyperparameters β.

Each text is determined by a conditional probability distribution $p(z|w_m)$ for a word, $p(z|\omega_m)$ which obeys a polynomial distribution with parameters θ_m, θ_m. The parameters obey a Dirichlet distribution with hyperparameters α_k.

Each word in each text is determined by the topic distribution $p(z|w_m)$ for that text, and the topic distribution $p(w|z_k)$ for all words.

(2) Model construction

First generate K a word distribution for the topic, randomly generate a parameter vector according to the Dirichlet distribution $Dir(\beta)$, and randomly generate a parameter vector $\varphi_k,\varphi_k \sim Dir(\beta)$ as the word distribution for the topic $p(w|z_k)$, $w\in W$.

Then the topic distribution of M texts is randomly generated according to the Dirichlet distribution $Dir(\alpha)$ and a random parameter vector $\theta_m,\theta_m \sim Dir(\alpha)$ as the word distribution of the topic $p(z|w_m)$, $m=1,2,\cdots,M$.

Then M words are randomly generated N_m. The process is as follows:

1) First follow a polynomial distribution $Mult(\theta_m)$ - generate a topic $Z_{mn},Z_{mn} \sim Mult(\theta_m)$

2) Then follow the polynomial distribution $Mult(\varphi_{z_{mn}})$ - randomly generate a word $\omega_{mn},\omega_{mn} \sim Mult(\varphi_{\Sigma_\infty})$

(3) Parameter setting

In the LDA topic model text generation process, the number of text topics K is given.

(4) Model learning algorithm solution

LDA topic model belongs to the learning belongs to unsupervised learning, it is difficult to solve accurately, this paper uses Gibbs sampling algorithm to approximate the solution. The basic principle is that by integrating the hidden variables θ and φ, the marginal function distribution $p(w,z|\alpha,\beta)$ is obtained, in which w observable variables z are unobservable; Gibbs sampling is performed on the posterior probability distribution $p(z|w,\alpha,\beta)$ to get distribution $p(z|w,\alpha,\beta)$ sample ensemble; this sample ensemble is then used to estimate the θ, φ parameters, and ultimately all of the LDA subject model $p(w,z,\theta,\varphi|\alpha,\beta)$. Parameter Estimation.

III. Analysis of the impact of Chinese feminism on the change of women's social status in modern China

III. A. Theoretical Value of Studying Feminism in the Perspective of Historical Materialism

In present-day China, the study of Western Marxist feminism is undoubtedly of great value in terms of how to better adhere to Mann's idea of women's emancipation and how to solve the complex problems faced by women in China at the present time. In theory and practice for China's women's emancipation cause has an important reference role, for the emancipation of women's thinking in the new situation to provide a wealth of theoretical materials, and constantly promote the development of Marxist women's emancipation thinking, so that it is more in line with the current situation, and a better guide to the practice of activities. Then we need to widely accept good sources, and only in this way can we stand in the forest of academics.

III. B. Analysis of high-frequency words

Guided by the materialistic view of history and the basic principle of “social existence determines social consciousness”, this study systematically integrates multi-dimensional historical materials and academic resources to construct a material system covering theoretical traceability, historical practice and reality. The 300 texts on modern women's issues between 1919 and 1949 were selected for word frequency statistics, and the top 10 high-frequency words were extracted after eliminating the deactivated words, and the results are shown in Table 1. The most frequent word is “women's emancipation”, reaching 9876.

Table 1: High-frequency Words

Rank	Frequency	Word
1	9876	Women's liberation
2	8789	Freedom of marriage
3	7721	Gender equality
4	6664	Social status
5	5602	Family
6	5531	Revolution

7	4506	Education
8	3433	Occupation
9	2401	Country
10	2365	Law

III. C. Thematic analysis based on LDA modeling

III. C. 1) Determining the optimal number of topics

According to the research purpose and research design, obtaining the optimal solution for the number of topics is the key to this study. Regarding the determination of the optimal number of topics, this paper adopts the perplexity degree as the index for evaluating the optimal solution for the number of topics, and selects the number of topics corresponding to the smallest perplexity degree as the number of topic classes for the final division of the research text. The numerator part of the logarithmic function of Perplexity is the negative of the likelihood estimation of generating the whole set of documents (which represents the generating ability of the parameters trained in the training set), and as the probability takes the value range of [0, 1], we know from the nature of the logarithmic function that the numerator value is positive and positively correlated with the text generation ability. The denominator represents is the total number of words in the document collection. That is, the stronger the model generating ability, the smaller the confusion value. The maximum number of topics is set to 15 and the number of iterations is set to 1000, the Perplexity of the given document set is calculated and the relationship between the number of topics and the perplexity is plotted as shown in Figure 4. From the previous analysis, it can be seen that the lower value of perplexity represents the more fitted the model is, and the value of topics selected according to the fold plot should also be the largest, but it is clear that when there are too many topics, the model has been overfitted. In the figure, the perplexity decreases as the value of K increases. When K belongs to [1, 12], the curve drops sharply. When K is greater than 12, the curve basically levels off. Therefore, the inflection point 12 is the optimal value of K. Therefore, the number of LDA topics generated in the selected corpus set of this paper is selected as 12.

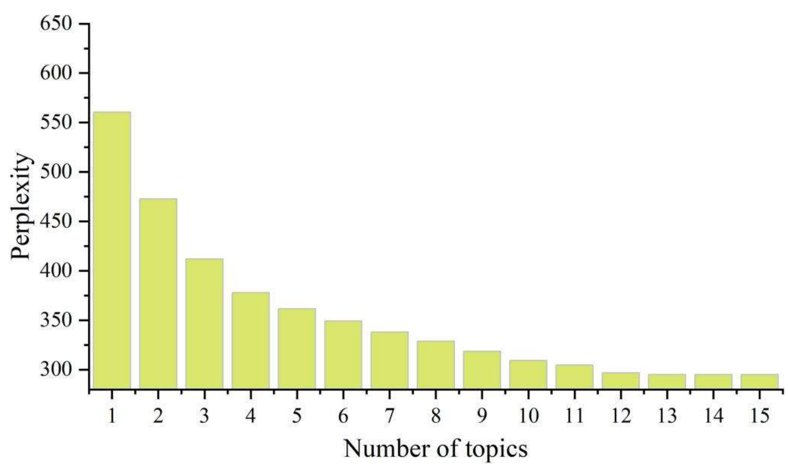


Figure 4: Relationship between the number of topics and the degree of confusion

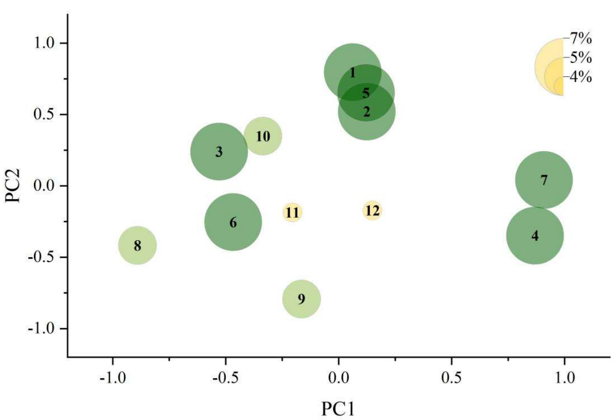
III. C. 2) Thematic analysis

Since the analysis of each theme in this paper needs to be combined with the view map to specify, so this paper will be automatically generated by the view map in the order of naming the theme 1-12. The order of the view map is automatically generated by the number of articles contained in each theme from high to low, and the weight of each theme in the document is calculated, the view map theme weights are shown in Table 2. The unit of measurement of the research sample in this paper is "article" as the standard, that is, each article will correspond to a theme. Table 2 visualizes the discussion focuses of modern women's issues from 1919-1949 and the social concerns in the historical context. Among them, the theme of "freedom of marriage" ranks first with a weight of 12.62%, significantly higher than other themes, which is highly consistent with the historical background of modern Chinese society against feudalism. The next most important theme, "revolution and gender liberation", reveals the deeply nested relationship between the revolutionary movement and women's liberation. Against the backdrop of the combination of Marxist ideas of women's emancipation and Chinese revolutionary practice, women are integrated into the overall narrative of class liberation, and their social advancement is no longer confined to the level of individual rights, but is closely related to the social revolution that overthrew the old system and established a new order.

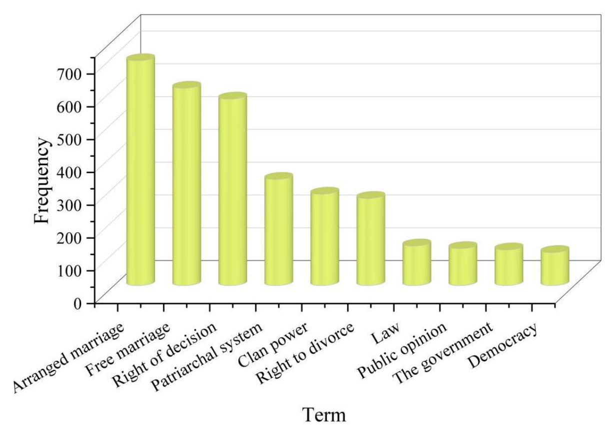
Table 2: Subject Weights of the line-of-sight map

Serial number	Topic	Label	Weight
1	Topic#5	Marriage Freedom	12.62%
2	Topic#8	Revolution and Gender Liberation	11.09%
3	Topic#3	Educational equity	10.38%
4	Topic#4	Workplace safety	9.04%
5	Topic#9	Family power	8.75%
6	Topic#11	Religion	8.62%
7	Topic#12	National development	8.58%
8	Topic#2	Human rights protection	8.52%
9	Topic#6	Women and Children's affairs	7.13%
10	Topic#7	International issues	6.02%
11	Topic#1	Campus management	4.74%
12	Topic#10	Protection of minors	4.51%

In this paper, we will conduct an in-depth analysis of the topic of "freedom to marry" in combination with the sight distance chart, and the sight distance map and the top 10 most relevant terms are shown in Figure 5 (a~b), respectively. It can be seen from the visual distance diagram that the theme of "freedom to marry" is close to the themes of "revolution and gender liberation" and "family power", reflecting that the three do not exist in isolation. The demand for freedom of marriage not only points to the struggle for the autonomy of individual marriage and love, but also involves the challenge of the traditional family system and social power structure. From the top 10 most relevant terms in Figure 5 (b), high-frequency words such as "arranged marriage", "free marriage", and "decision-making power" constitute the core semantic field of the topic. Terms such as "patriarchy", "clan rights", and "right to divorce" reveal the institutional roots of the impediments to the freedom of marriage.



(a) Intertopic Distance Map



(b) Top 10 Most Relewant Terms

Figure 5: Visualization of the theme of "Marriage Freedom"

III. D. Analysis of the living conditions of women's families in recent times

This paper combines historical document analysis and quantitative empirical methods to systematically examine the status change of women in the family field in modern China. The experimental group is selected from the group of women who have received new-style education or participated in women's liberation organizations after 1919, with a total of 100 cases. The control group selected traditional family women who had not received new-style education or participated in women's organizations during the same period. Stratified sampling was matched by geography, urban and rural areas, and ethnicity, and an effective sample of 100 cases in the experimental group and 100 cases in the control group was finally determined.

III. D. 1) Marital autonomy

The results of the distribution of initial marital decision-making power among women with different levels of education in the experimental and control groups are shown in Table 3. The distribution of first-time marriage decision-making between the experimental group and the control group by education level shows a significant gradient difference. Specifically, in the experimental group, the proportion of women deciding to marry on their own shows an increasing trend as the education level increases. In the illiterate group, only 60% of the females were able to decide on marriage independently, which rose to 92.86% in the elementary school group, and the proportion of females in the junior high school group and in the secondary school and above education level reached 100%. In the control group, the proportion of autonomous decision-making is significantly lower than that of the experimental group in all education level groups. This result shows that new-style education directly enhances women's subjectivity in marriage decision-making by breaking the cognitive constraints imposed on women by traditional rituals, and the higher the level of education, the more significant this facilitating effect is.

Table 3: Distribution of women's first marriage decision

			The parents decided to seek their own opinions	Decided by myself	I have decided to seek my parents' opinions	Total
Experimental group (N=100)	Educational attainment	Illiteracy	8	0	12	20
		Primary school	2	0	26	28
		Junior high school	0	0	38	38
		Secondary vocational school or above	0	6%	8	14
	Total		10	6	84	100
Control group (N=100)	Educational attainment	Illiteracy	16	0	6	22
		Primary school	6	0	20	26
		Junior high school	12	0	28	40
		Secondary vocational school or above	6	0	6	12
	Total		40	0	60	100

The actual residence status of the experimental and control groups after marriage is shown in Table 4. The proportion of the experimental group who set up their own households after marriage is significantly higher (46%) than that of the control group (40%), and the proportion of those who return to their husbands' households is lower (36%) than that of the control group (48%). The traditional family system of “living with one's husband” has been gradually dissolved in the newly educated female group, and replaced by a housing pattern that is more in line with modern individual consciousness.

Table 4: Actual residential status after marriage by gender

	The man's family	The woman's family	Independent portal	Total
Experimental group (N=100)	36	18	46	100
Control group (N=100)	48	12	40	100

III. D. 2) Decision-making on important family matters

Decision making in major economic affairs of the family in the experimental and control groups by education level is shown in Table 5. In the experimental group, the proportion of joint decision-making by husband and wife in major family economic affairs is significantly higher than that of the control group. For example, in the case of “purchasing major items”, the proportion of joint decision-making by husband and wife in the experimental group is 68%, which is 38 percentage points higher than that in the control group. In the case of “investment or loan”, the proportion of joint decision-making in the experimental group was 66%, compared with 24% in the control group. The new style of education not only enhances women's awareness of economic participation, but also promotes the transformation of the family's decision-making model into one of equality.

Table 5: Decision-making Situation of Major Family Economic Affairs

			Illiteracy	Primary school	Junior high school	Secondary vocational school or above
Experimental group (N=100)	Purchase large items	Husband	12	10	4	0
		Wife	0	2	4	0
		Husband and wife	8	16	30	14
	Investment or loan	Husband	10	12	2	2
		Wife	2	2	4	0
		Husband and wife	8	14	32	12
Control group (N=100)	Purchase large items	Husband	20	20	20	4
		Wife	0	0	4%	2%
		Husband and wife	2	6	16	6
	Investment or loan	Husband	22	20	22	6
		Wife	0	2	4	0
		Husband and wife	0	4	14	6

III. E. Multiple regression analysis

In this paper, overall social status satisfaction and conformity to social status satisfaction were used as outcome variables, respectively, and years of education, occupational participation, legal awareness, annual family income, and age were used as antecedent variables, and stepwise regression analysis was conducted using SPSS software. The results of the multiple regression analysis are shown in Table 6. Comparing the coefficient of determination R^2 and the adjusted coefficient of determination AR^2 of the two models, it is found that the coefficient of determination and the adjusted coefficient of determination of the composite social satisfaction measurement model are 0.42 and 0.48, respectively, whereas the coefficient of determination and the adjusted coefficient of determination of the overall social satisfaction measurement model are 0.16 and 0.24, which indicates that the decomposition measure of social status is superior to the overall measure. Therefore, the following analysis will be based on the output of the composite social satisfaction measurement model. Years of education, occupational participation, legal awareness and annual household income have a significant positive effect on women's social status satisfaction. Among them, the influence of years of education is the most prominent (β=0.32), and the improvement of education level directly enhances women's cognition and recognition of their social status. The standardized coefficient of occupational participation (β=0.28) is the next most important, and the significant effect of legal awareness score (β=0.25) indicates that the awakening of rights awareness is an important psychological mechanism for the improvement of women's social status. Annual household income has a positive but weak effect (β=0.18), while age fails the significance test. Overall, the model explains about 48% of the variance in social status satisfaction, which is highly compatible with the basic principles of historical materialism. Changes in the socio-economic structure and political system of modern China ultimately contributed to the advancement of women's social status by altering their educational, occupational and legal practices.

Table 6: Results of multiple regression Analysis

	Overall social satisfaction			Compound social satisfaction		
	Standardized regression coefficient	T value	Sig.	Standardized regression coefficient	T value	Sig.
Years of education	0.29	3.89	<0.001	0.32	4.12	<0.001
Professional participation	0.25	3.21	<0.001	0.28	3.56	<0.001
Legal cognition	0.22	2.87	<0.01	0.25	3.12	<0.01
Annual household income	0.16	2.01	<0.05	0.18	2.23	<0.05
Age	-0.07	-1.05	>0.05	-0.08	-1.23	>0.05
R^2	0.42			0.48		
AR^2	0.16			0.24		
F value	18.67			20.23		

IV. Conclusion

Based on the materialistic view of history, this study synthesizes the methods of text mining and empirical analysis to systematically explore the impact of Chinese feminism on the change of women's social status in modern times, and draws the following main conclusions.

Among the 300 texts on women's issues in modern times, the most frequent word is “female liberation”, reaching 9876, and the theme of “freedom of marriage” ranks first with a weight of 12.62%, which is significantly higher than that of other themes.

In the experimental group, the proportion of women deciding on their own about marriage shows an increasing trend as the level of education rises. Only 60% of women in the illiterate group were able to decide on their own about marriage, rising to 92.86% in the elementary school group, and reaching 100% in the middle school group and in the secondary school and higher education levels. In the control group, the proportion of autonomous decision-making is significantly lower than that of the experimental group in all education levels. The proportion of self-supporting households after marriage in the experimental group (46%) is significantly higher than that in the control group (40%), and the proportion of returning to the husband's family (36%) is lower than that in the control group (48%). At the same time, the proportion of joint decision-making by husband and wife in major economic affairs of the family was significantly higher in the experimental group than in the control group.

Years of education, occupational participation, legal knowledge and annual household income have a significant positive effect on women's social status satisfaction. Among them, the influence of years of education is the most prominent (β=0.32), followed by the standardized coefficient of occupational participation (β=0.28), and the significant influence of legal cognition score (β=0.25) suggests that the awakening of the awareness of the rights is an important psychological mechanism for the improvement of women's social status. Annual household income has a positive but weak effect (β=0.18), while age fails the significance test. Overall, the model explains about 48% of the variance in social status satisfaction, which is highly compatible with the basic tenets of historical materialism.

References

[1] Evans, S. M. (2015). Women's liberation: Seeing the revolution clearly. Feminist Studies, 41(1), 138-149.
[2] Oksala, J. (2018). Feminism, capitalism, and ecology. Hypatia, 33(2), 216-234.
[3] Haug, F. (2016). Marxism-feminism. Historical Materialism, 24(4), 257-270.
[4] Federici, S. (2018). Marx and feminism. TripleC: Communication, Capitalism & Critique. Open Access Journal for a Global Sustainable Information Society, 16(2), 468-475.
[5] Barclay, K. (2017). New Materialism and the New History of Emotions. Emotions: History, Culture, Society, 1(1), 161-183.
[6] Armstrong, E. (2020). Marxist and socialist feminisms. Companion to feminist studies, 35-52.
[7] Cruz, K. (2018). Beyond liberalism: Marxist feminism, migrant sex work, and labour unfreedom. Feminist Legal Studies, 26(1), 65-92.
[8] Bohrer, A. (2018). Intersectionality and Marxism: A critical historiography. Historical Materialism, 26(2), 46-74.
[9] Marx, U. (2019). Accounting for equality: Gender budgeting and moderate feminism. Gender, Work & Organization, 26(8), 1176-1190.
[10] Cozza, M., & Gherardi, S. (2023). Feminism under erasure in new feminist materialism as a case of symbolic manspreading. In A research agenda for organization studies, feminisms and new materialisms (pp. 33-54). Edward Elgar Publishing.
[11] Sun, J. Y., & Zhuang, J. Y. (2022). Family role and social status. In Chinese women in leadership (pp. 17-34). Cham: Springer International Publishing.
[12] Hossain, D. M., Ahmad, N. N. N., & Siraj, S. A. (2016). Marxist feminist perspective of corporate gender disclosures. Asian Journal of Accounting and Governance, 7, 11-24.
[13] Swirsky, J. M., & Angelone, D. J. (2016). Equality, empowerment, and choice: what does feminism mean to contemporary women?. Journal of Gender Studies, 25(4), 445-460.
[14] Lykke, N. (2020). Transversal dialogues on intersectionality, socialist feminism and epistemologies of ignorance. NORA-Nordic Journal of Feminist and Gender Research, 28(3), 197-210.
[15] Schild, V. (2019). Feminisms, the environment and capitalism: on the necessary ecological dimension of a critical Latin American feminism. Journal of International Women’s Studies, 20(6), 23-43.
[16] Wilson, K. (2015). Towards a radical re-appropriation: Gender, development and neoliberal feminism. Development and change, 46(4), 803-832.
[17] Casalini, B. (2017). A materialist analysis of contemporary feminist movements. Anthropological Theory, 17(4), 497-517.
[18] Kessler, J. K. (2025). Law and historical materialism. Duke LJ, 74, 1523.