Hindawi
Complexity
Volume 2024, Article ID 9815987, 1 page
https://doi.org/10.1155/2024/9815987

*Retraction*

# Retracted: Forecast and Simulation of the Public Opinion on the Public Policy Based on the Markov Model

**Complexity**

Received 23 January 2024; Accepted 23 January 2024; Published 24 January 2024



This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

## References

[1] Z. Li, "Forecast and Simulation of the Public Opinion on the Public Policy Based on the Markov Model," *Complexity*, vol. 2021, Article ID 9936965, 11 pages, 2021.

Hindawi
Complexity
Volume 2021, Article ID 9936965, 11 pages
https://doi.org/10.1155/2021/9936965

*Research Article*

# Forecast and Simulation of the Public Opinion on the Public Policy Based on the Markov Model

**Zi Li** [1,2]

[1]*National Security Institute, People's Public Security University of China, Beijing 100038, China*
[2]*Department of Investigation, Shandong Police College, Jinan, Shandong 250000, China*

Correspondence should be addressed to Zi Li; lizi@sdpc.edu.cn

Received 25 March 2021; Revised 23 April 2021; Accepted 7 May 2021; Published 19 May 2021

Academic Editor: Zhihan Lv



Public policy and public opinion directly affect the image of the government, but due to the lack of appropriate monitoring and early warning tools, the government's handling of credit changes is seriously lagging behind. In response to this problem, this paper integrates the internet, public information, market credit information, and other data, uses hidden Markov models and natural language processing technology, and establishes a modern government public policy and public opinion monitoring and early warning model to evaluate government credit in real time; the government can formulate relevant policies based on the evaluation results to improve the government's governance capabilities. Empirical analysis shows that, based on the dynamic scoring framework and Markov model, the government credit monitoring and early warning models established, respectively, have 90% of the reference value, and the analysis results have the same reference. This method can effectively predict the trend of hot online public opinion. The subsequent establishment of an online public opinion early warning system and an online public opinion guidance mechanism provided theoretical support.



## 1. Introduction

Internet public opinion (IPO) refers to the sum of emotions, attitudes, and opinions expressed by individuals or various social groups and organizations on various public things that care about themselves or are closely related to their own interests through online channels [1]. Internet public opinion is the mapping of the social network public opinion in the internet space, and it is a direct reflection of the social network public opinion. Policy decision-making and evaluation in the field of national public management is a complex and dynamic issue involving changes in environmental conditions. The effective implementation of public policies is the foundation of social operations and has a great impact on our daily lives. The implementation of new policies is often necessary to complete more spanning and challenging attempts in a short time. Improper or invalid policies not only affect our quality of life but also damage the credibility of the government [2]. The use of new tools and methods such as online public opinion to simulate national public policies will help us to use a more effective and accurate way to predict and study the relevant issues of policy implementation. It is the only way for the government to make scientific decisions [3]. In recent years, due to the popularity of the internet and the rapid development of social media (such as Weibo and MSN), netizens' online activity has increased day by day, and the influence of the online public opinion has increased day by day. The subsequent negative online public opinion is very likely to induce the unhealthy emotions of the people which lead to violations or excessive behaviours, which seriously affect social stability [4]. With the advent of the Web 2.0 era, the popularization of information technology has changed people's thoughts and behaviours and, at the same time, has brought new challenges to the management of social public affairs [5]. From the perspective of online public opinion, with the advancement of the democratization of the country's society, the public's awareness of participation in politics is increasing, and the cognitive tendency of internet users on current policies can feed back a lot of hidden

information, which promotes the formulation and evaluation of national public policies [6]. However, the current public policy forecasting research in the country mostly focuses on the exploration of the policy evaluation index system, the research on the evolution trend of public opinion, and the analysis of thematic viewpoints. Few articles have analysed the development of the policy public opinion [7]. Therefore, studying the communication characteristics and laws of the online public opinion will play an important role in the healthy development of an online culture and social environment.

Existing research techniques mainly include quantitative analysis using bibliometric methods, trend analysis of the network public opinion based on the Markov chain, application of uncertainty theory to deal with the uncertainty of the network public opinion, and emergent network public opinion evolution based on game models. The recognition and tracking based on the single-pass clustering model are studied, the SOM neural network clustering is analyzed, and the small-world network virus spreading based on the cellular automata model is studied [8]. In recent years, the national internet has developed rapidly. Both the network infrastructure and the number of netizens have developed greatly. According to the released statistical report on the development of the internet, as of the end of the last year, the total number of internet users in the country has reached 100 million people [9]. It can be seen that the popularity of the internet in the country is becoming more and more widespread. As a new communication medium, the internet has an unparalleled advantage over traditional media. Under the influence of the rapid development of this new media, social public opinion has appeared in a new form of online public opinion. Due to the difference in its communication media, it has the characteristics of convenient and fast expression, diversified information methods, and strong data fluidity. The information structure has fundamentally changed the past single "publish-accept" message architecture. Any one of these points may play multiple roles such as publishing, disseminating, modifying, and accepting. This is a complete deconstruction and complete subversion of the form of communication based on traditional media. The characteristics of virtuality, anonymity, diffusivity, permeability, arbitrariness, real time, etc., have a profound impact, resulting in a lot of differences in the content and form compared with the traditional public opinion due to the effect of policy implementation. It is closely related to the trend of various national livelihood indicators. In order to reasonably predict and evaluate the feasibility and accuracy of government decision-making, many scholars have used computer technology to conduct empirical simulation studies on multiple livelihood policy fields [10]. This paper integrates internet credit information and other data and uses the hidden Markov model to establish a modern government public policy and public opinion monitoring and early warning model, which can realize the evaluation of government credit.

This article first constructed a policy support measurement model, took China's upcoming public opinion policy as an example, and compiled public opinion texts on Weibo as the input data for policy simulations. Second, the public regards the change of policy attitude as a random process and uses the Markov model to simulate and predict its future direction. The experimental results can provide a decision-making basis for the further improvement and implementation of the policy. At the same time, combined with grounded theoretical methods, opinion mining, and other natural language processing technologies, it integrates a variety of theoretical models, focusing on the comprehensive application of qualitative and quantitative research, thereby making research possible. The conclusion has good explanatory power and maneuverability. Specifically, the model consists of two parts: in the first stage, the socialization parameters of the prediction module are generated through the judgment model supported by the policy and public opinion; in the second stage, the Markov theory is used to realize and improve the prediction model. Finally, the dynamic error compensation formula is used to modify the prediction accuracy. We seek the iterative balance point of the system and discuss the MATLAB simulation results.

## 2. Related Work

At present, scholars and experts in various fields have begun to pay attention to the public opinion, network public opinion, and public opinion crisis, and related research reports and academic literature are also increasing. Scholars have conducted related literature and paper research related to social sciences, humanities, natural sciences, and other disciplines. Among them, for the research on "internet public opinion," the earliest research on journal documents started in 1984, and the research on dissertations started in 1984. On the whole, the related research on the internet public opinion is still a relatively new research topic. As scholars pay more and more attention to it, its development is very rapid.

At present, more popular public policy-public opinion simulation prediction methods can be divided into three types: CGE general equilibrium model, SD system dynamics model, and ABM main body modeling model. Since the popularity of the online public opinion is a random process and is discrete with respect to time and state, a Markov model can be established to analyse and predict the development trend of the popularity of the online public opinion.

The prediction of the public policy-public opinion under the general equilibrium model is based on the balance of supply and demand in economics. It lists a number of equilibrium equations by observing the resource structure or operating relationship in real cases and seeks for certain conditions. The best results of the principal component utility of each variable have been shown. At present, most foreign research examples focus on economic and trade policies, using econometric models to establish demand forecasting frameworks or to improve economic accounting systems in a hierarchical manner. Some Western countries represented by the United States have begun to build economic policy guidance simulation systems (Fair model, Murphy model, MSG2 model, etc.) and have achieved initial

results. These systems are used to guide the formulation, implementation, and evaluation of economic policies for the people's livelihood. Then, they make the government have an absolute initiative in the policy process. The system dynamics model was first proposed by Zhang et al. [11]. It takes structure, control, feedback, and function as the four key points. It establishes a dynamic model through a combination of qualitative and quantitative reasoning and then simulates and predicts the system through computer simulation. Subsequently, Roney et al. [12] first used this model to conduct simulation experiments on urban development policies through several defining variables such as population, land, and employment and evaluated the practical effects of such policies. At present, the scope of applied research based on this model is the most extensive at home and abroad. For example, Pankowska et al. [13] built a national medical policy system based on this model and proposed that the integration and evaluation of several system variables can effectively promote the efficiency of related policy decisions. Some researchers selected carbon dioxide emissions as the simulation result parameter to select and optimize environmental policies [14]; Boeschoten et al. [15] simulated five subsystems of economy, market, policy, environment, and application industries, targeting the coal-bed methane industry. The policy puts forward a number of recommendations; Bertoni et al. [16] simulated the impact of several variables on the national food supply and demand situation and provided certain theoretical support for the national oil and grain policy analysis. The so-called subject modeling research is based on the observation of event reality and simulates the dynamics and complexity of the system through the simulation interaction between a number of significant norms of behavior subjects. This model skips the stage of theoretical framework construction and is known as the closest to the society. Recently, the ASPEN system developed by Enns and Wohlfarth [17] based on the subject model can be regarded as a classic example in the field of economic policy simulation. On this basis, the giant ABM framework proposed by Huang et al. [18] in 2004 has created a new framework for policy simulation. There are also many policy simulation practices that use ABM modeling in China. By comparing various policy simulation research methods, it is concluded that the agent model is more suitable for evaluation in the process of financial policy formulation; Zhou [19] used swarm as a simulation environment to build a financial balance and the theoretical ABM model and proposed that this kind of policy simulation result that makes full use of statistical data is more time efficient and scientific. Aiming at the defects of the hidden Markov model (HMM) in the cross-site scripting detection of the inaccurate estimation of the initial a priori hypothesis and the poor classification ability of the HMM parameters specified by the maximum likelihood criterion, a cross-site-based MLP-HMM is proposed. First, natural language processing (NLP) is used to solve the problem of high-dimensional complexity of the data. Then, the weights of the entire model are fine-tuned through multilayer perceptron (MLP) neural network learning to obtain the initial observation matrix.

Regarding the application of web crawlers in public policy-public opinion prediction, earlier related research can be traced back to the F-S model, which is an improvement and optimization of the model. Later, search models based on certain specific areas and topic-based crawlers also gradually appeared. Most of the poster-based crawling strategies are based on this framework. Since then, more and more research results, such as context-based crawlers, have proposed models for marking seven communities and proposed evaluation indicators for topic crawlers. In recent years, well-known search engine companies at home and abroad have implemented many key research technologies and successfully applied them to the internet. In order to overcome the limitation that the model is suitable for data attribute clustering, the KS model came into being. The limited iterations of the model can only converge to the local minimum and cannot achieve the global optimum. The clustering model based on the grid density contours applies the iterative initial point set refinement model to the model.

Regarding the design of the text clustering method in the public policy-public opinion model, it is mainly divided into divided text clustering and hierarchical text clustering. The direction of partitioned text clustering can be traced back to the mean clustering model proposed in 2019. Regarding the research of hotspot acquisition, foreign countries have proposed topic discovery methods based on language models. Dai and Serletis [20] proposed a method for calculating word weight based on the log-likelihood test and proposed a method for identifying subtopics by adding named entities, etc., and proposed a method for identifying hot topics with free parameters. The research on forecasting originated from the autoregressive model established in 2011 to predict the law of market changes. It was originally based on the moving average autoregressive mixed model. Years later, with the development of signal processing, various models have been optimized and improved in theory. At the same time, the concept of grey system prediction is proposed, which solves the problem of learning the connection weights of hidden units in multilayer networks. The Markov chain was first proposed by Markov. Kolmogorov's theory of probability created conditions for its development. A data calculation method based on probability statistics proposed in the mid-1940s is often used in finance and the fields of engineering and computational physics.

## 3. Public Opinion System Framework of the Public Policy Based on the Markov Model

*3.1. Principle of the Markov Model of the Optimal Path.* Markov model believes that, for a certain type of discrete event, if there are several changing states and each state is only connected with the previous state after being arranged in a time series, then the process meets a certain probability conversion condition during the change [21–23]. It can be solved through a mathematical matrix to predict the unknown state. Subsequently, improvements and research studies on Markov forecasting models emerge in endlessly. Compared with other forecasting models such as grey forecasting, regression analysis, and neural network, this

model has lower requirements for data [24, 25]. Even if the stability of the experimental data is poor, it can still obtain relatively ideal results within the controllable error range. It has been applied to forecasts in the fields of economy (stock market, corporate indicators, and housing prices), trade (sales, transportation, and markets), environment (climate and pollution emissions), traffic (flow and accidents), etc. The benchmark test does not simply output pass/fail. The result of each test is a scalar, which means that we cannot simply fold the passed results. We can take a look at the data chart, perhaps we can have an intuitive understanding of the data pattern; after all, under normal circumstances, the number of benchmark tests is far less than that of correctness tests. Although we have been trying to produce stable and consistent results in benchmark tests, the curve will still vary greatly, depending on the size of the workload and the equipment being run. For example, compared to other workload benchmark data, we found that the test results are very unstable, and setting the threshold to one percent does not achieve the desired results in every test.

There are three convertible states ($S1$, $S2$, and $S3$). The time series of each state is represented by $t = 1, 2, 3, ..., n$, and the corresponding change result is $qt$. $t$ always satisfies

$$U\{v_t = s3|v(t-1) = s3, v(t-2) = s2\}, \tag{1}$$

that is, the state at time $S3$ is only related to time $S2$ and has nothing to do with time $S1$, so the process is called a Markov process. When extended to other multistate processes, it can be expressed as

$$U\{x(t_n) > x(t_n)\}|x(t_m - 1) = v(t_m - 1), x(t_m - 2) = (t_m - 2). \tag{2}$$

Markov theory has three important properties: no aftereffect, characteristics of the Chapman–Kolmogorov equation, and ergodicity. Any predictive model based on this theory is based on these three. $\lambda = [X, \pi, A]$ is used to represent the Markov model, where $X$ represents the $n$ observable states of the event, that is, $X = (x1, x2, x3, \ldots, x-n)$; $\pi$ represents the initial probability distribution in different states, that is, $\pi = \{\pi - i\ \}$, $i = 1, 2, \ldots, n$; $A = P$ (which is the matrix formed by the transition probabilities of each state, $P - i - j = P(x - j|x - i)$ which means that at $t - 1$, the state is transformed from $xi$ to the probability of $x - j$. Then, the above three properties can be expressed as follows. Property 1: the state probability transition has no aftereffect. For any sequence $\{u1, u2, \ldots, u - t\}$, there are

$$\left[u(t+1)\,\middle|\,u1, u2, \ldots, u_n\right] = \left[u(t+1)|u_n\right], \tag{3}$$

that is, the state at time $t + 1$ is only affected by the state at time $t$ and has nothing to do with other states. This property is a necessary and sufficient condition for the Markov model. For example, the probability of whether there will be a sunny day tomorrow is only related to today's weather conditions and has nothing to do with the weather conditions before today. Property 2: Chapman–Kolmogorov equation properties, also known as stationary properties: let $P(k) = (P - i - j(K))$ be the $K$-step transition probability of the Markov chain. For any positive integer $m$, there is

$$U_{st}^{(m+n)} = \sum_{n=1}^{m} U_{sn}^{m} U_{nt}^{m}, \tag{4}$$

Further extension shows that $P(k) = Pk(1)$, and the $k$-step transition matrix can be derived from the one-step transition matrix $P(1)$. Property 3: the ergodicity of the Markov process: for the Markov process $\{X(t),\ t \geq 0\}$, there exists $\forall j \in S$, and the probability after $k$ times of state transitions tends to a certain limit:

$$\begin{gathered} \lim_{x \longrightarrow \infty} U_{st}^{m+n} = \theta_t, \\ \lim_{s \longrightarrow \infty} U(v(m+1)) = \lim_{s \longrightarrow \infty} U(v(m)) U^t(s) = C, \end{gathered} \tag{5}$$

that is, for a stochastic process system, starting from any state, if it is not affected by major qualitative changes, it will become stable when it develops to a certain stage and obtained the state transition matrix. The state transition matrix is composed of the transition probabilities of each state. It is known that $\{X(t),\ t \in T\}$ is used to represent the Markov process, the state space is $S = \{S1, S2, \ldots, Sn\}$, the state $Si \in S$ at each moment, $Si$ has $n$ kinds of transformation possibilities: $Si \longrightarrow S1$, $Si \longrightarrow S2$, $\ldots$, $Si \longrightarrow Sn$, and if the conversion is completed in one step, the conversion probability is expressed by the formula, namely, $\forall s,\ t \in S$:

$$U_{st}(m) = U(v(m+1) = t|v(m) = s). \tag{6}$$

The matrix set composed of $S$-step conversion probabilities is the $S$-step probability conversion matrix, which can be expressed as

$$U_{st}^{m+n} = \begin{bmatrix} U_{11}^{m+n} & U_{12}^{m+n} & U_{1s}^{m+n} \\ U_{21}^{m+n} & U_{22}^{m+n} & U_{2s}^{m+n} \\ U_{s1}^{m+n} & U_{s2}^{m+n} & U_{ss}^{m+n} \end{bmatrix}. \tag{7}$$

From the stationarity of the Markov model mentioned above, it can be deduced that the $k$-step transition matrix and the one-step transition matrix have a power multiplication relationship, that is, it satisfies

$$\begin{cases} U_{st}(m) \geq 0, & s, t \in R, \\ \sum_{i=1}^{k} U_{st}(m) = 1, & s, t \in R. \end{cases} \tag{8}$$

The original data mostly appear in mutually independent forms. The state space can be divided according to the actual observation value and membership function. For example, the behavior of coin toss can be divided into two states according to the observed positive and negative values; rainfall can be divided according to the amount. The level of subordination is divided into three states: light rain, moderate rain, and heavy rain. According to this, it can be assumed that the random event $\{X(t),\ t \in T\}$ has $m$ discrete values, which can be divided into $n$ state spaces $S = \{S1, S2, \ldots, Sn\}$. The initial probability value is determined. The

initial transition probability is the probability of dividing the observation value into each state space for the first time. Here, only the first $m-1$ observation points are considered. If there are $Ci$ values divided into state $Si$, then $\sum Ci = m-1$, then $Si$ appears the probability of $Pi = Di = Ci\ m-1$. Transition probability matrix $P$ is solved in one step. The state space $Si$ contains $Ci$ values. Assuming that a total of $C-i-j$ observation points are included in the state $S-j$ when the transition occurs, the transition probability is $P-i-j = D-i-j = C-i-j$, $Ci$, and the matrix composed of $P-i-j$ is the one-step state transition matrix for sequence inspection. When the transition frequency is known, the Markov property of the sequence can be checked by 2. $f-i-j$ is the quotient of the sum of the column frequencies of the frequency matrix and the total frequency. In the case of a certain level of significance, if 2 obeys the distribution of $(m-1)2$, then the data satisfy the Markov characteristic and predict the state of the sequence. According to the calculation probability of the state transition matrix, contact the previous state matrix to iteratively obtain the final prediction result. Here, the number of iterations can be set to calculate the desired time node, and the final prediction value can also be selected through the equilibrium state.

*3.2. Prediction and Evaluation Process of the Public Policy and Public Opinion System.* The construction of the public policy-public opinion system model is mainly divided into three stages. (1) Data preparation: it mainly completes the collection and preprocessing of the Weibo comment corpus. (2) Mining public attitudes and opinions: the purpose is to calculate the public sentiment ratio (positive, medium, and negative) of the policy in each dimension. The HMM state transition process is based on the Markov random field (MRF), that is, the independent distribution of the probability of each state. It can be derived from the formula that the probability of $q$ is only related to its previous state. This stage is divided into two steps: first, for each microblog, use the domain frame semantic dictionary and policy review topic word dictionary to extract the evaluation words and topic words in the text. On this basis, the matching relationship between the frame and the topic is used as a template to identify each policy evaluation dimension of Weibo. Second, an improved sentiment tendency analysis algorithm is used to calculate the sentiment value of each microblog and construct a policy evaluation vector. (3) Dimensional sentiment synthesis: the policy evaluation vectors of all microblogs are used to count the public's sentiment tendency in each policy evaluation dimension, and the AHP level analysis and DF document frequency, two weight measurement methods, are used to synthesize the public's public opinion support for the policy. The characteristic layer of the specific public policy-public opinion system is shown in Figure 1.

In order to transform the original public policy-public opinion corpus into structured evaluation data, it is necessary to construct a policy evaluation vector for the analysis of the support of the policy public opinion. The policy evaluation vector is defined as $E-Si = (ESi1,\ ESi2,\ ESi3,\ ESi4)$ (3–6). Among them, $ESi$ is the evaluation vector of Weibo; $ESi1$, $ESi2$, $ESi3$, and $ESi4$, respectively, represent the sentiment values on the four policy dimensions. The artificial neural network method does not need to make prior distribution assumptions about the input signal distribution, has strong self-organization and self-learning capabilities and good classification effects, and can make up for the defects of the HMM. Therefore, this paper adopts the method of combining the multilayer perceptron neural network and HMM to realize anomaly detection. The public opinion tendency under the policy dimension can be obtained through the policy evaluation vector. According to its numerical value, Weibo is divided into positive ($E-Si > 0$), medium ($E-Si = 0$), and negative ($E-Si < 0$). We get three emotion levels, and count the proportion of samples under each emotion. The synthesis of the overall public opinion support of the policy takes the $L1$, $L2$, $L3$, and $L4$ dimensions as the target-level evaluation indicators and weights. In general, the common methods for determining indicator weights are divided into subjective assignment method and objective assignment method. This article intends to adopt analytic hierarchy and DF document frequency, comprehensively considering the factual connection between the subjective opinions of experts and data samples to explore the two types of weight of the degree of influence of the value on the experimental results.

The emotional tendency in public policy-public opinion is not only determined by the evaluation word itself but also affected by the degree adverbs and negative words. For example, "the working people on the frontline strongly oppose it." Although the word "strong" does not change the polarity of the semantics itself, it increases the intensity of negative emotions. In order to improve the accuracy of emotion value calculations, this article combines the review text with the Chinese grammar dictionary and selects 37 degree adverbs $D-k$ ($k = 1, 2, \ldots, 37$) and defines the emotional strength of each word $\mathrm{Sen}(D-k)$. To a certain extent, the negative word assumes the role of the emotional valence movement indicator (VSI), so the sentence sentiment tends to be expanded or reversed, but the negative words in some idioms often do not affect the polarity of the sentence. The specific algorithm is shown in Figure 2. Among them, img represents the mean square error, imgread represents the value of the $i$-th output of the MLP when the vector $0f$ is used as the input, transform represents the target value, $T$ is the length of the input sequence, and $M$ is the number of output nodes. If it is too large, the convergence speed will cause instability, and if it is too small, the convergence speed will be slower. By adding the momentum term, the adjustment speed is increased, and the oscillation is reduced during stable adjustment. Among them, the BP initial learning rate is set to 0.001; decrease rate indicates the proportion of each attenuation, set to 0.99; it indicates how many steps to decay at a time, and the parameter is set to 50. In this paper, combining the original negative words in the How-Net dictionary and Weibo common phrases, a total of 27 negative words $N-m$ ($m = 1, 2, \ldots, n$) are sorted out manually, and the formula is used to calculate the negative

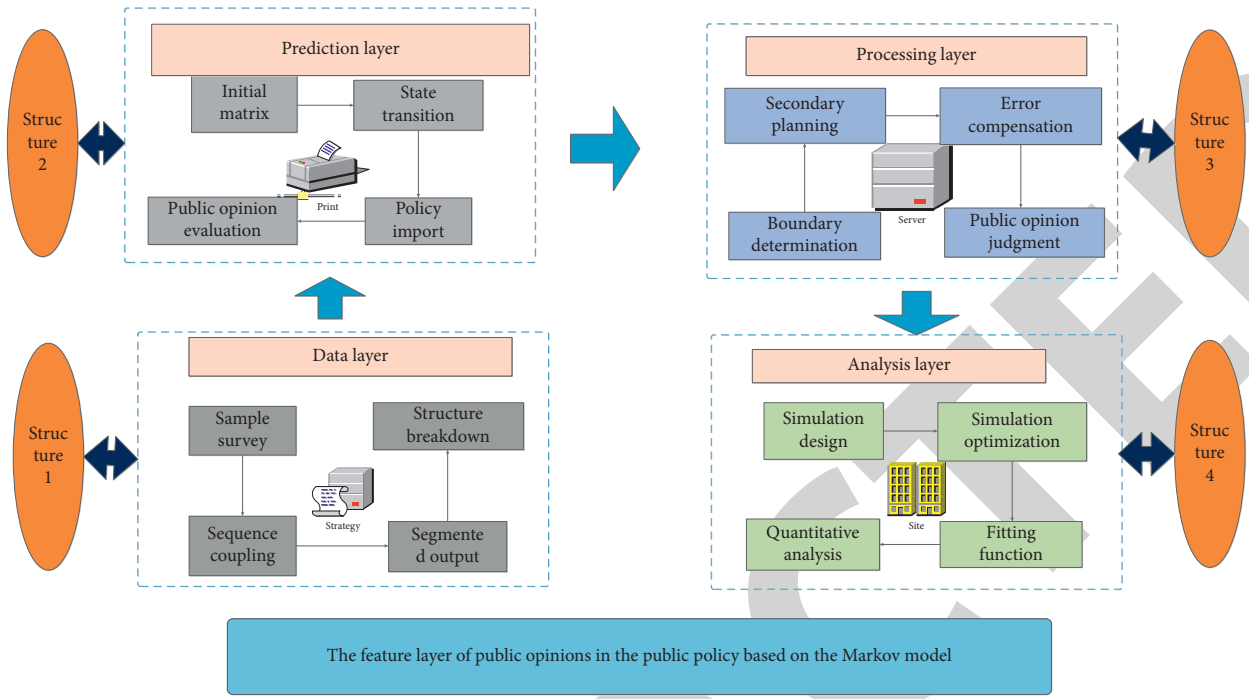


Figure 1: The feature layer of public opinions in the public policy based on the Markov model.

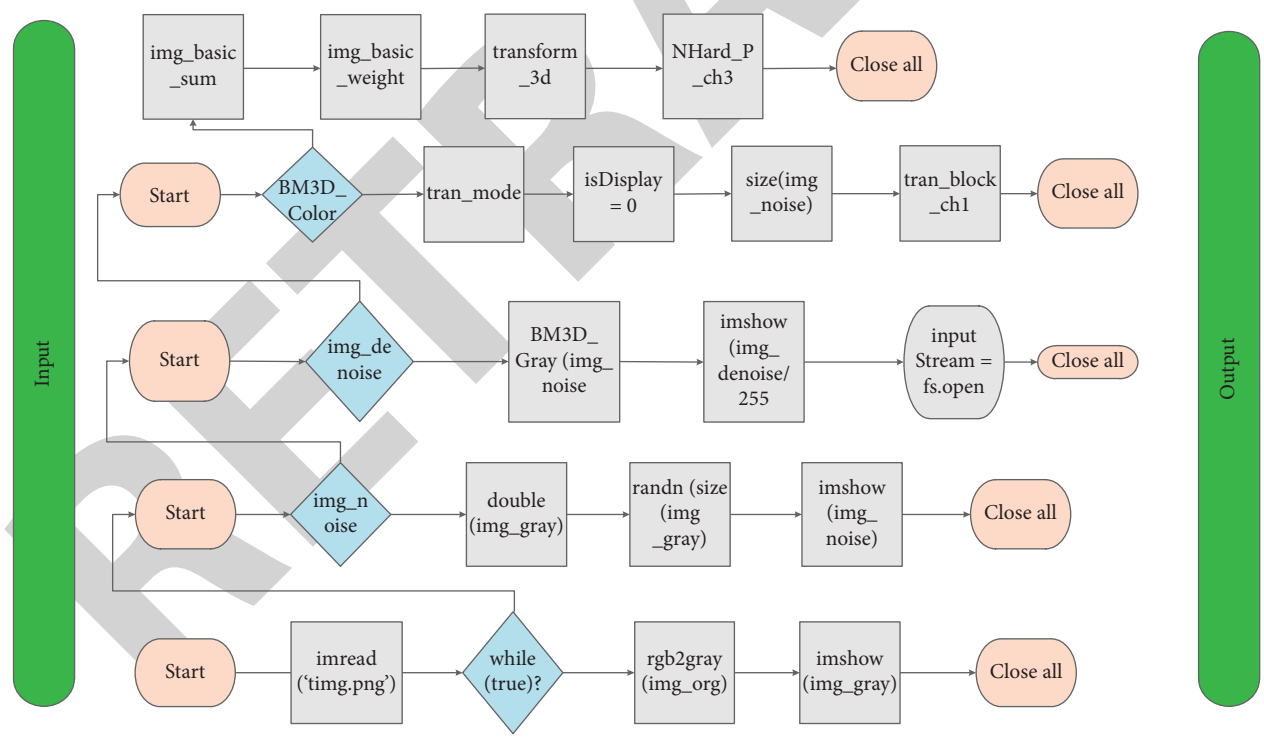


Figure 2: Public policy-public opinion system forecast algorithm diagram.

words $N-m$ in each Weibo; the degree of influence $\mathrm{Inf}(N-m)$ determines its emotional polarity. At present, the research on national text public opinion recognition is relatively scarce. Here, the initial sentiment value $E'(Si)$ of the comment is mainly used to judge the superstate state to find whether the comment is ironic. The specific method is as follows: with the help of the concept of numerical measurement of the degree of intermediary truth value, $E'(Si)$ is sorted in a directional order, and the "superpositive" or "supernegative" individuals in the data sample are found by setting emotional boundaries. Rule R1: rhetorical sentences with the $E(Si)$ value in the superpositive range,

regardless of whether there is continuous punctuation in the sentence or at the end of the sentence, are regarded as derogatory sentences. Rule R2: exclamation sentences with the $E(Si)$ value in the superpositive range, When odd negative words and multiple punctuation marks appear at the same time, they are regarded as derogatory sentences.

*3.3. Optimization of Public Opinion Weights of the Public Policy Based on the Markov Model.* From the perspective of policy, this paper analyzes the public's attitude and tendency through the combination of the policy evaluation dimension identification framework and the semantics of the framework and the dictionary and establishes the analysis model of public opinion support based on social media. The construction of this model is mainly divided into three major frameworks, as shown in Figure 3. (1) Data preparation: NLP obtains the posterior probability $P$ through training, that is, the input layer vector $x$. The posterior probability is used as the initial observation probability of the HMM, and the state transition probability obtained by $N$-gram processing is substituted for the HMM to calculate the node probability, and the node probability is less than the threshold $p$ as error. Here, when the sliding window of the threshold is 2, the minimum node probability of the optimal path in the sequence is obtained. The HMM optimized with MLP not only improves the detection rate of the original HMM in XSS detection but also shortens the detection time. Compared with traditional machine learning algorithms, the model proposed in this article also has certain advantages in cross-site scripting detection. The dataset structure obtained by preprocessing the XSS dataset is shown below, and three sets of mixed datasets $D1$, $D2$, and $D3$ are selected, and the structure is shown below. The parameter settings of the hybrid model are also updated at the same time.

Through the identification of text evaluation words and topics, it can lay the foundation for the following comment policy dimension identification and sentiment analysis. Due to the sparseness of the data, the experimental results obtained by the dictionary-based lexical weight accumulation algorithm are not ideal. The reason is mostly due to the use of sentence structures such as public opinion and satire in Weibo, which leads to the sentiment value judged by the machine not necessarily reflecting the sentimental tendency of the comment. Since the internet public opinion uses some exaggerated rhetoric to express satire and dissatisfaction with social events and its emotional value is significant, this article uses the hyperstate thought in MMTD to identify the public opinion sentences in the experiment and combines the punctuation feature rules to determine the emotional polarity of comments. Using the Viterbi algorithm to obtain the optimal state sequence can be divided into three basic steps: starting from the first state, for its $1n$ nodes, calculate their probability to represent any state 1 node. Because there is only one step, these distances are the maximum probabilities from $S$ to their respective ones. For all nodes in the second state, calculate the maximum probability from $S$ to them. The weight distribution of each node is shown in Figure 4. For a specific node, the path from $S$ to it can pass through any node in $1n$ of state 1, and the corresponding path length can be understood as the weighted sum of probabilities. Next, follow the second step method from the second state to the third state, until the last state, to get the optimal path, that is, the optimal state sequence of the text. The state sequence with the largest output probability represents the optimal state sequence implied by the observation sequence, and then the semantic orientation of the text can be analysed through the optimal state sequence. The design is based on the sentiment analysis model of HMM, and the sentiment status set is (neutral, oppose, agree). Under the given HMM, the observation sequence $T$ can be generated by the following steps: choose an initial state $S-q1$ according to the probability $i$ of the initial state; set $t1$; According to the output probability distribution of the state $k-bi\ \ k-v$; is output the state transition probability distribution $i-j-a$, transfers the state at the current time $t$ to the new state $j-t1$; if $T-t$, repeat 3; otherwise, end the calculation.

The weight distribution in the public opinion model is mainly composed of two parts: the judgment module of policy public opinion support based on online comments and the public opinion prediction module based on Markov theory. Among them, the first module mainly uses natural language processing technology that combines frame semantics and emotional dictionaries to explore the current public attitude trends of a certain policy, as shown in the following text. To put it simply, a single text evaluation vector is constructed through the two main lines of the network comment text and the current policy evaluation dimensions, and then sentiment analysis technology is used to calculate the sentiment value, and finally, the final public opinion tendency is integrated according to the weight of each dimension. The second module is the simulation and prediction of future public opinion trends. This process is based on Markov random theory and uses the user's social influence to update the emotional data of the first module as the input parameters of the model, build a MATLAB simulation environment, and use genetic algorithms, and the quadratic optimization programming and dynamic error compensation formula are used to solve the experiment.

## 4. Application and Analysis of the Public Opinion Based on the Markov Model

*4.1. Mapping Simulation of the Observation State and Hidden State of the Public Opinion Model.* In the setting of the hidden Markov status, first, determine the direction of the credit change of the public policy and public opinion subject, which is generally given according to the direction of public opinion change. If there are three states of public opinion tendencies, namely, "positive," "neutral," and "negative," then the government's credit status can be given three states in practical applications, namely, "up," "maintain," and "down." If there are two states of emotional orientation, namely, "positive" and "negative," after determining the observed state and the hidden state, the mapping relationship and state transition probability need to be determined. For the convenience of calculation in the following

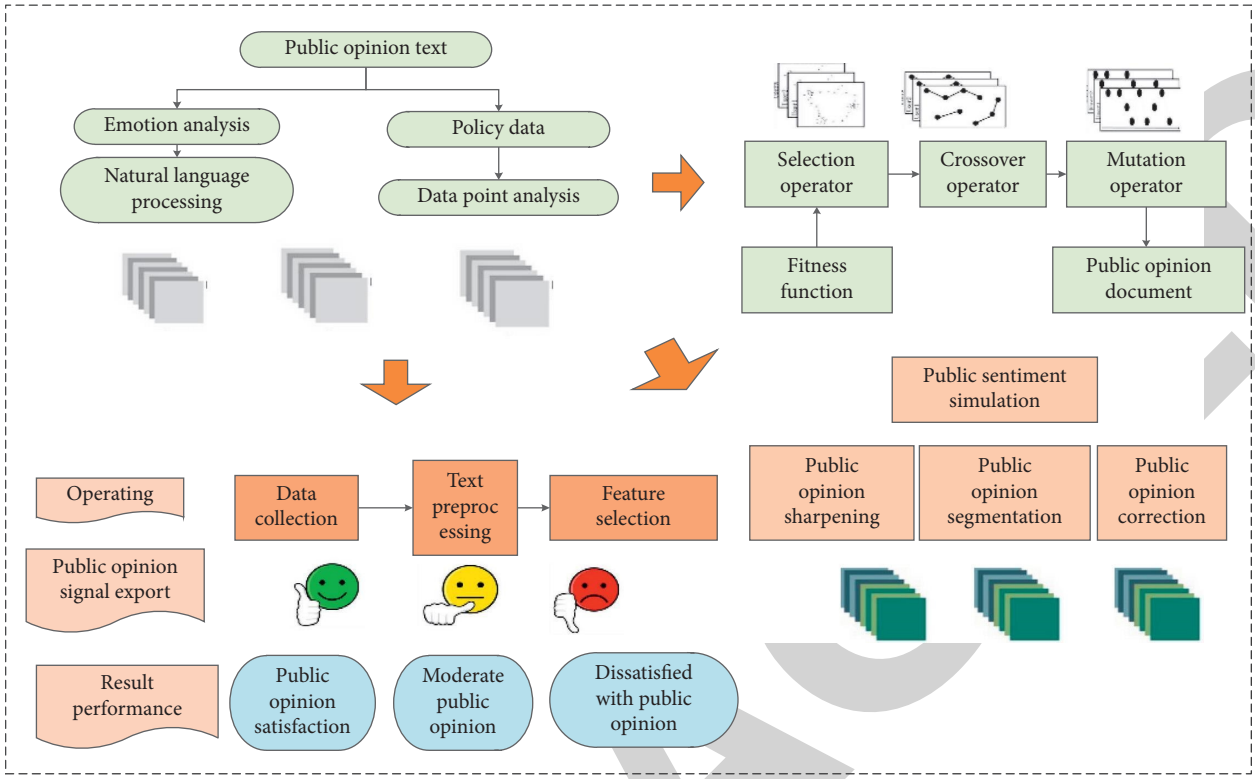


Figure 3: The framework of the public policy and public opinion system based on the Markov model.

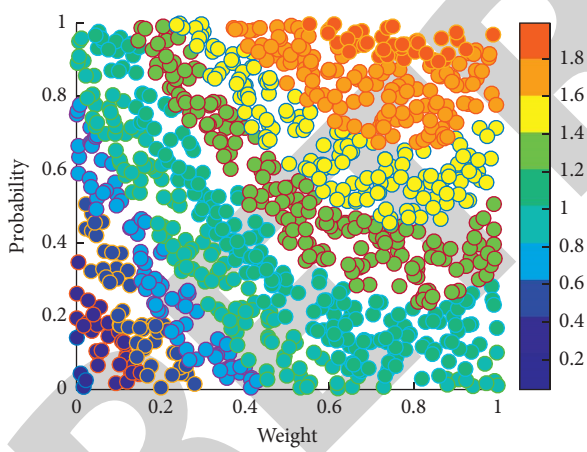


Figure 4: Distribution of node weights of the public policy and public opinion data.

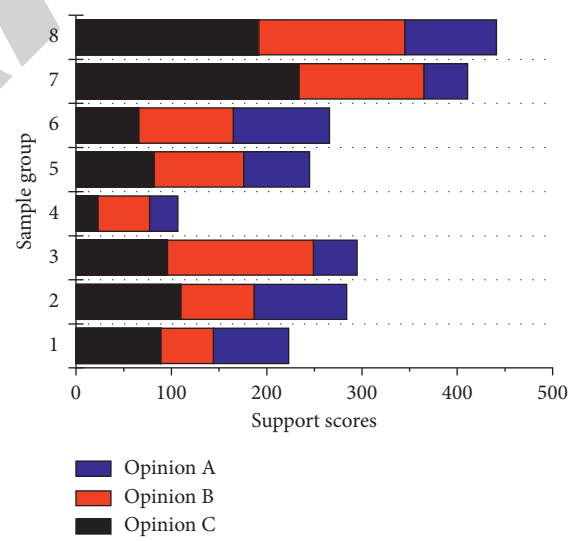


Figure 5: Distribution of public policy and public opinion support.

examples, here, we take two types of implicit states of the subject, "upregulation" and "downregulation," as examples and set the observation state of the subject, that is, three public opinion states: "positive," "neutral," and "negative"; the corresponding support distribution is shown in Figure 5.

The experimental data in this paper are evaluated using the public opinion text corpus provided by COAE2014, and a total of 500 texts in various fields are selected as the dataset. For all the documents to be evaluated, the word segmentation NLPIR/ICTCLAS2015 tool is used to perform word segmentation, part-of-speech tagging, and preprocessing of useless words, and then a public sentiment test is conducted. This paper selects the social media platform as the data source platform, through three stages of data crawling, data cleaning, and preprocessing, to obtain online comments on public opinion policies from the public as the corpus data for public opinion support analysis: (1) data crawling: use the open-source software "Octopus" big data collector, set the cycle login collection and delayed loading rules with the selected time range as the interval, collect the original comment text on the public opinion in a time sequence, and

collect more than 12,000 related comments. The collected content includes nine fields including user name, comment text, user follow number, user posted number, number of fans, number of comments, number of likes, and number of reports. (2) Data cleaning: between Weibo comments subjective and arbitrary, there are no unified expression standard and different expression forms. Therefore, it is necessary to preprocess the data to eliminate irrelevant elements and clarify the emotional factors. The purpose is to restore the natural language as much as possible and minimize noise in the emotional processing stage. This is mainly to complete the deletion of URL links ("http://url"), tag links ("#hash-tag"), username ("@username"), and the replacement of emoticons and traditional characters. (3) Preprocessing: in this paper, the ICTCLA developed by the Chinese Academy of Sciences is used to segment the comments and part-of-speech tagging, and the data are structured to lay the foundation for the following text sentiment analysis. At the same time, manual correction is performed to reduce experimental errors. The obtained data before and after simulation are shown in Figure 6.

With the help of the frame dictionary of the public opinion domain, the frame word elements of the text can be extracted, which can be used to match the policy dimension table and calculate the sentiment value. Among them, lemma $W - j$ is the vocabulary with a specific meaning in the field of policy evaluation, and a total of 123 words are screened with the help of the network policy corpus. The emotional strength $W - j$ of the word element is represented by a numerical value, less than 0 indicates negative emotion, 0 is based on the empirical prediction and analysis of public opinion policies, and greater than 0 is positive emotion. The framework is a lexical classification organization. After sorting and partial adjustments, a total of 12 framework roles have been formed, namely, emotional experience, acceptance, fairness, difficulty, time, means, strength, flexibility, clarity degree, effectiveness, desirability, and magnitude. The design of the index weight is based on the analytic hierarchy process (AHP) to verify its consistency, that is, the consistency test. Test formula: $CR = CI/RI$, where CI is the consistency index of the comparison matrix and RI is the average random consistency index. When $CR < 0.1$, the consistency check passes, and the decision can be made according to the result represented by the vector; when $CR > 0.1$, the check fails, indicating that there is a contradiction in the judgment process, and it should be adjusted or restructured. The consistency test shows that $CR = 0$, and the consistency test is passed. This paper also uses AHP to test the consistency of other dimension indicators, and all of them have passed the test. After the satisfaction survey, it can be seen that the public policy and opinion system based on the Markov model performs well and is supported by the vast majority. The specific parts are as shown in Figure 7. The criteria for determining the indicators are based on the way of keyword extraction and comparison. With the help of program judgment, it is judged whether the corresponding index is satisfied or not, and finally, the evaluation result is generated according to the judgment result.

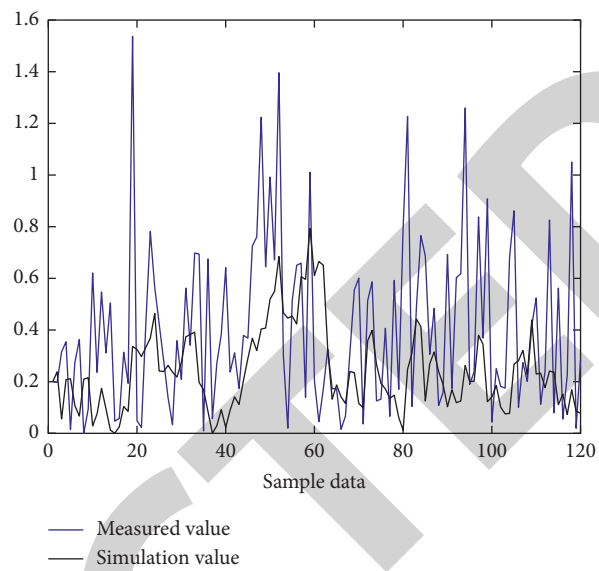


Figure 6: Comparison of predictions of the public opinion before and after simulation.

*4.2. Example Results and Analysis.* Obtain public policy and public opinion-related data from relevant mainstream websites, combine relevant information to filter out irrelevant and repetitive events, and obtain useful data as examples of this research. The collected data mainly come from major websites, communities, forums, etc. Search for keywords on Baidu News, Baidu webpages, Sogou webpages, and Tianya forums to obtain data. The data released by the event are summarized in units of days and numbered in order to get the amount of data released every day in the time span. According to the previous formula, calculate the popularity of the online public opinion due to uncertain factors in order to exclude trend fluctuations caused by abnormal points, and select tail data for prediction. We select the initial data as the time of data collection to test the accuracy of the model. The state vector corresponding to the trend value of the network public opinion at the initial moment is $w$. Use the initial state vector and state transition probability matrix, predict the probability of the trend value of the network public opinion in each state after 5 days, compare the predicted value with the real data of event development, and perform error comparison analysis.

It can be seen from Figure 8 that the combined prediction method based on the grey model and Markov chain has achieved better results than the original single top measurement method, and each error index is lower than the single prediction error index. The example initially verifies the combined prediction, the rationality, and the feasibility of the model. Combining with the "nonobjection" concept of the public opinion of the policy audience mentioned in the previous article, the research results of this article are shown in Figure 9. It can be seen that, in the initial stage, with step 3 as the dividing line, the public's nonobjection increased and opposition decreased, followed by a reversal phenomenon. The nonobjection declines, and the opposition ratio rises. Until step 17, the public's nonobjection ratio for the deferred

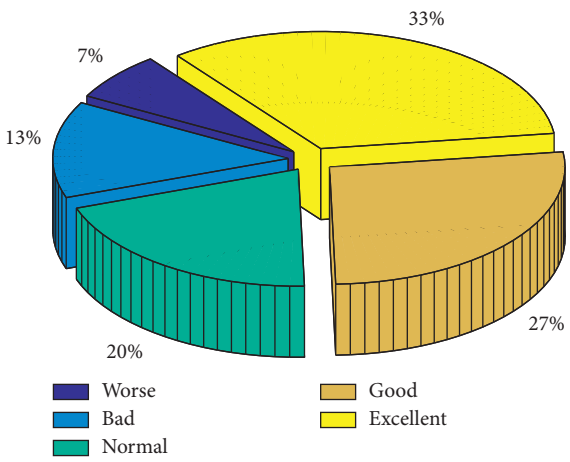


Figure 7: Satisfaction survey of the public policy and public opinion system based on the Markov model.

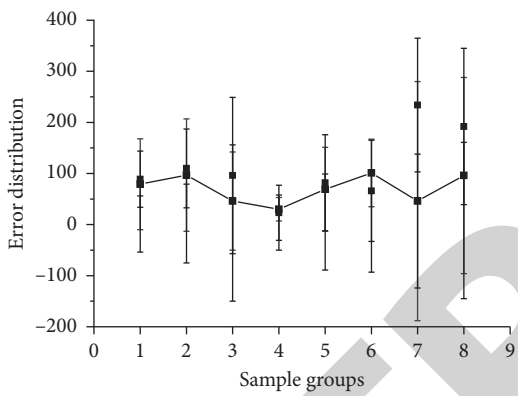


Figure 8: The distribution of public policy and public opinion forecast errors.

policy in a balanced state remains at about 50.58%, which is less than the feasibility standard value of 0.75, so its feasibility is open to discussion. Based on the score and the path of score change in 2017, the model can calculate the comprehensive evaluation score of the public policy-public opinion in a city in 2018 to be 76.749, while the score given by the National Information Center is 83.82. The consistency between the scoring results and the results of the National Information Center is 91.56%, indicating that the model has a certain validity.

The experimental results show that firstly, after comparing the actual policy trends, the prediction model proposed in this article can effectively count the people's online sentiment tendencies towards policies in the future; secondly, in the prediction stage, compared to the solution of a single genetic model, the combination proposed in the article models and compensation formulas can improve the accuracy of experimental predictions. Finally, the case analysis shows that the current public's nonobjection to the deferred policy is relatively low, and there are greater policy risks. The government needs to continuously coordinate various works in the policy evaluation beforehand to ensure

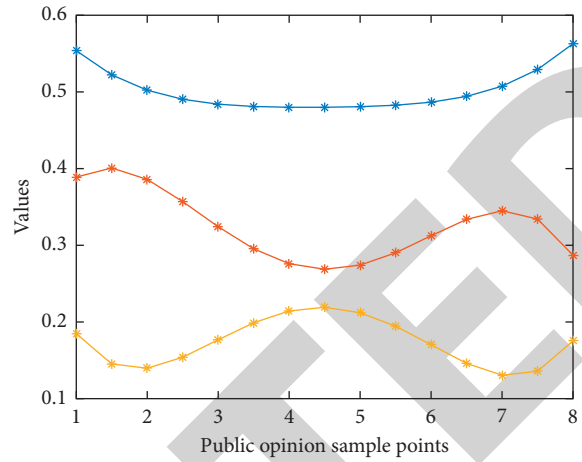


Figure 9: The distribution prediction fitting curve of different attitudes of the public policy and public opinion.

its smooth implementation. At the same time, in response to the current concept of the policy and the national social media public opinion supervision work, this article puts forward a number of suggestions, hoping to make some contributions to the research of public opinion in the policy field in the future.

## 5. Conclusion

When using the Markov forecasting model, this paper selects the method of combining genetics and quadratic programming optimization. The selection of the parameters of the heritage algorithm is mainly obtained through experiments, and the method of improving the fitness is not studied through the theoretical system. For the sentiment analysis improvement method in the public opinion support analysis model, the sentiment tendency of the internet public policy and public opinion is identified and modified mainly in accordance with the actual situation. Since most of the related research on natural language mining has used the standard corpus as the experimental data, focusing on improving the efficiency and accuracy of the experiment, the theoretical basis of the public opinion recognition and improvement methods proposed in this paper is feasible. In follow-up research, attention should be paid to the development of policy public opinion research, combined with machine learning methods, to expand more application contexts for public opinion comments, and to improve analysis efficiency.

## Data Availability

The data used to support the findings of this study are available upon request to the author.

## Conflicts of Interest

The author declares that there are no conflicts of interes.

## References


[1] X. Liu, T. Tang, and D. He, "Double-layer network negative public opinion information propagation modeling based on continuous-time markov chain," *The Computer Journal*, vol. 2, no. 8, pp. 39–47, 2020.

[2] X. Hong, C. Zhang, W. Huang, and W. Hong, "Prediction of food Safety network of public opinion hot-degree based on improved entropy method and markov," in *Fuzzy Systems and Data Mining IV*, vol. 309, pp. 30–43, IOS Press, Amsterdam, Netherlands, 2018.

[3] D. Riedlinger and J. Wisniewski, "On mass killing in america: using markov analysis to challenge conventional understanding & policy design," in *Proceedings of the A Regional Conference of the Society for Industrial and Systems Engineering*, New York, NY, USA, May 2018.

[4] E. A. Markov, "The origins and traditions of the information policy in russia," *Comparative Politics*, vol. 10, no. 1, pp. 70–80, 2019.

[5] D. Naskar, E. Onaindia, M. Rebollo, and S. Das, "Modelling emotion dynamics on twitter via hidden markov model," in *Proceedings of the 21st International Conference on Information Integration and Web-Based Applications & Services*, pp. 245–249, Munich, Germany, 2019.

[6] S. Vermeer and D. Trilling, "Toward a better understanding of news user journeys: a markov chain approach," *Journalism Studies*, vol. 21, no. 7, pp. 879–894, 2020.

[7] Q. Chen, X. Guo, and H. Bai, "Semantic-based topic detection using Markov decision processes," *Neurocomputing*, vol. 242, pp. 40–50, 2017.

[8] S. Liu, J. Teng, and Y. Gong, "Extraction method and integration framework for perception features of public opinion in transportation," *Sustainability*, vol. 13, no. 1, 254 pages, 2021.

[9] C. Chino and A. Acconcia, *Markov Discrete Choice Process for Dividend Policy*, University of Naples Federico II, Naples, Italy, 2019.

[10] N. El-Mawass, P. Honeine, and L. Vercouter, "SimilCatch: enhanced social spammers detection on twitter using markov random fields," *Information Processing & Management*, vol. 57, no. 6, Article ID 102317, 2020.

[11] L. Zhang, T. Wang, Z. Jin, N. Su, C. Zhao, and Y. He, "The research on social networks public opinion propagation influence models and its controllability," *China Communications*, vol. 15, no. 7, pp. 98–110, 2018.

[12] J. P. Roney, J. Ferlic, F. Michor, and T. O. McDonald, "ESTIpop: a computational tool to simulate and estimate parameters for continuous-time markov branching processes," *Bioinformatics*, vol. 36, no. 15, pp. 4372-4373, 2020.

[13] P. Pankowska, B. F. M. Bakker, D. L. Oberski, and D. Pavlopoulos, "How linkage error affects hidden markov model estimates: a sensitivity analysis," *Journal of Survey Statistics and Methodology*, vol. 8, no. 3, pp. 483–512, 2020.

[14] Y. Lan, Z. Lian, R. Zeng et al., "A statistical model of the impact of online rumors on the information quantity of online public opinion," *Physica A: Statistical Mechanics and Its Applications*, vol. 541, Article ID 123623, 2020.

[15] L. Boeschoten, D. Filipponi, and R. Varriale, "Combining multiple imputation and hidden markov modeling to obtain consistent estimates of employment status," *Journal of Survey Statistics and Methodology*, vol. 7, pp. 43–57, 2020.

[16] D. Bertoni, G. Aletti, G. Ferrandi, A. Micheletti, D. Cavicchioli, and R. Pretolani, "Farmland use transitions after the CAP greening: a preliminary analysis using Markov chains approach," *Land Use Policy*, vol. 79, pp. 789–800, 2018.

[17] P. K. Enns and P. C. Wohlfarth, "Making sense of the supreme court-public opinion relationship 1," *Routledge Handbook of Judicial Behavior*, pp. 180–195, 2017.

[18] X. Huang, H. Jin, and Y. Zhang, "Risk assessment of earthquake network public opinion based on global search BP neural network," *PLoS One*, vol. 14, no. 3, Article ID e0212839, 2019.

[19] X. Zhou, "Hierarchical item response models for analyzing public opinion," *Political Analysis*, vol. 27, no. 4, pp. 481–502, 2019.

[20] W. Dai and A. Serletis, "On the Markov switching welfare cost of inflation," *Journal of Economic Dynamics and Control*, vol. 108, Article ID 103748, 2019.

[21] L. Xiong, H. Zhang, Y. Li et al., "Improved stability and H∞ performance for neutral systems with uncertain markovian jump," *Nonlinear Analysis: Hybrid Systems*, vol. 19, pp. 13–25, 2016.

[22] B. Wang, F. C. Zou, J. Cheng et al., "Fault detection filter design for continuous-time nonlinear markovian jump systems with mode-dependent delay and time-varying transition probabilities," *Advances in Difference Equations*, vol. 2017, no. 1, pp. 1–23, 2017.

[23] J. Yang, J. Zhang, and H. Wang, "Urban traffic control in software defined Internet of things via a multi-agent deep reinforcement learning approach," *IEEE Transactions on Intelligent Transportation Systems*, pp. 1–13, 2020.

[24] X. Liu, H. Song, and A. Liu, "Intelligent UAVs trajectory optimization from space-time for data collection in social networks," *IEEE Transactions on Network Science and Engineering*1 page, 2020.

[25] J. Wen, J. Yang, B. Jiang, H. Song, and H. Wang, "Big data driven marine environment information forecasting: a time series prediction network," *IEEE Transactions on Fuzzy Systems*, vol. 29, no. 1, pp. 4–18, 2021.