

Retraction

Retracted: Research on Prediction of News Public Opinion Guiding Power Based on Neural Network

Security and Communication Networks

Received 11 July 2023; Accepted 11 July 2023; Published 12 July 2023

Copyright © 2023 Security and Communication Networks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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) Peer-review manipulation

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] S. Chen and G. Li, "Research on Prediction of News Public Opinion Guiding Power Based on Neural Network," *Security and Communication Networks*, vol. 2022, Article ID 2607492, 9 pages, 2022.

Research Article

Research on Prediction of News Public Opinion Guiding Power Based on Neural Network

Shan Chen  and **Guang Li**

School of Media Communications, Linyi University, Linyi, Shandong 276000, China

Correspondence should be addressed to Shan Chen; chenshan@lyu.edu.cn

Received 18 February 2022; Revised 7 March 2022; Accepted 22 March 2022; Published 18 April 2022

Academic Editor: Chin-Ling Chen

Copyright © 2022 Shan Chen and Guang Li. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The development of big data technology and the popularity of we-media platforms make the prediction of news public opinion more complicated, which means that the complexity and dynamic nature of news public opinion require higher accuracy of prediction. With the rapid popularization of we media technology, traditional single-model algorithm is difficult to effectively predict network public opinion under the current background. Therefore, this paper proposes an algorithm based IGA-RBF neural network to deal with the complicated news public opinion prediction. Firstly, the ARMA (autoregressive moving average model) prediction model is constructed and the BRN neural network is combined. Then IGA is introduced to optimize BRN neural network, and the column vector of output matrix of hidden layer is optimized globally. The algorithm uses k-means clustering to select parameters in RBF network. The experimental results demonstrate that the model algorithm makes up for the shortcomings of the single prediction algorithm, improves the accuracy of prediction, and has better prediction results of public opinion trends.

1. Introduction

With the rapid development of Internet technology, China has entered a new era of information explosion, and users scattered on the network have become the creators of information [1]. Thus, the speed and breadth of information dissemination have spread exponentially with the booming development of online social media. While facilitating information dissemination, social individuals also add complexity to public opinion monitoring and prediction [2]. Internet and we-media users tend to be younger people who prefer to obtain news information through Weibo, WeChat, Douyin, and other means [3]. The information content released by traditional media is gradually ignored by people, and also, the authoritative information is less and less. The amount of information released through the Internet is so large that it is difficult to distinguish authenticity. News public opinion without legal constraints is prone to public opinion crisis [4]. Intense language caused by public opinion in the society tends to have a negative impact on the values and cognition of netizens, and negative news and public

opinion increase the difficulty of social management [5]. When news time harms the image of the government, it will have a very negative effect on our government. Therefore, the study of the evolution of news public opinion is of great significance to the development of society [6]. New technologies and new platforms are used to quickly find news hotspots, grasp the development trend of news public opinion through prediction algorithms, and correctly guide news public opinion comments, so as to promote the healthy development of social and cultural industries [7].

With the continuous expansion of the scale of Internet users, the Internet has become another important way of information dissemination besides traditional media such as newspaper, radio, and TV [8]. At present, social media or websites relying on Internet platforms such as WeChat, Tiktok, Weibo, forums, Post bar, and news websites have become important battlefields of information dissemination [9]. An increasing number of netizens receive information from the real world and the virtual world and express their views and opinions on social events freely on social media. When sudden network events or social events attract a large

number of netizens' attention, they evolve into online public opinions, which will have a significant impact on social public security and long-term development [10]. In the era of big data, a huge amount of data needs to be collected for information work. Meanwhile, social media, news media, search engines, and other network platforms make network communication present a "honeycomb" divergent structure, making it more difficult to analyze and predict network and situation [11]. In addition, the warning time is unstable with the outbreak of emotional events. These problems make the response to online public opinions face greater challenges.

Literature [12] first clearly defined the negative public opinions of the government by taking microblog as the information carrier. Then it establishes the government negative public opinion prediction star based on Markov chain, which provides theoretical support for the government to deal with negative public opinion timely and reasonably and guide the trend of public opinion. Lyapunov was used in literature [13] to strongly prove that the trend development of network has chaotic characteristics, and related data were reconstructed in phase space to prove the practicability of the algorithm. Literature [14] improved the accuracy of microblog public opinion prediction by combining improved SEIR model and PageRank algorithm with Bayesian network. Literature [15] proposed an emotion perception time series prediction method based on dynamic time warping and autoregressive comprehensive moving average model to predict topic popularity. Literature [16] innovatively constructed Baidu Index time series index of network events and adopted relevant time data for training. Literature [17] studies the prediction and discovery of Weibo hot topics from the perspectives of data mining and term terms, respectively. Literature [18] studies the changing trend of public emotion based on the emotional common sense of microblog time. Literature [19] analyzes the mechanism and trend prediction of Wang Min's mood changes under the background of public opinion big data.

Many algorithms still need to adjust a large number of parameters in the optimization process, which increases the complexity of adjusting parameters in the process of network training, resulting in more network iterations and slower convergence. Aiming at the above problems, this paper proposes a research algorithm of news public opinion prediction based on improved RBF neural network. The innovations and contributions of this paper are listed below.

- (1) The algorithm uses k-means clustering to select parameters in RBF network.
- (2) AR-RBF combined prediction model is constructed, and historical data are selected as training samples of RBF neural network and ARMA time series.
- (3) The normalized order of RBF neural network was optimized by IGA. On the premise of satisfying the target error, the column vector of the output matrix of the hidden layer of the network is optimized.
- (4) The crossover probability and mutation probability of genetic calculation are adjusted adaptively according to the current individual fitness and the

number of evolutionary iterations, aiming to find the optimal orthogonalization order, reduce the average error, and improve the accuracy of prediction.

The chapter structure of this paper is as follows. The related theory is described in Section 2. The recommendation method is constructed in Section 3. Section 4 focuses on the experiment and analysis. Then Section 5 is the conclusion.

2. The Related Theory

2.1. RBF Neural Network. Radial basis function (RBF) [20] neural network does not fall into local extremum and can approximate any nonlinear function with arbitrary precision. Therefore, RBF neural network is chosen as the algorithm of public opinion prediction. RBF neural network is composed of input layer, hidden layer, and output layer. Its structure is shown in Figure 1.

In this paper, the commonly used Gauss function is chosen as the radial basis function of RBF neural network.

$$R_y(I) = \exp\left(-\frac{\|I - C_y\|^2}{2\sigma_y^2}\right), \quad (1)$$

where y is the number of nodes of the hidden layer. $R_y(I)$ is the output of the node at the hidden layer y of the network. I is the input vector. C_y is the center of the y radial basis function. σ_y is the scale factor of the y implicit node, which determines the radial action width of the function. Then through continuous learning, radial basis algorithm is used to adjust the weight ω_{yw} from the y implicit node to the w output node to minimize the learning error. Finally, the connection weight is fixed and the corresponding output is obtained from the input of the network.

2.2. Parameter Selection Based on K-Means Clustering. Because the number of hidden layer nodes y of RBF neural network is too small, it will lead to high error. Increasing the number of hidden layer nodes y can reduce the error of training. But too much increasing will affect the generalization ability of neural network and increase the network complexity, resulting in slow training speed. In principle, the minimum number of hidden layer nodes should be selected to meet the accuracy requirements. The key parameters of hidden layer are center C_y and width σ_y of hidden layer node. The center of traditional RBF neural network is the sample. Its width is determined by the number of data centers and the maximum distance between them and cannot be dynamically optimized. Therefore, it is necessary to determine the proper center C_y and width σ_y of the hidden layer to improve the performance of RBF neural network.

K-means clustering algorithm is a common unsupervised clustering algorithm with fast training speed and good ductility. Therefore, k -means clustering algorithm is used to determine the center C_y and width σ_y of the hidden layer to optimize the RBF neural network. The initial category of k -means clustering algorithm samples is unknown, so the

number of clustering centers needs to be set. In RBF neural network, the center C of Gaussian function is the clustering center. After determining the initial cluster center, formula (2) is used to calculate the distance between the input layer data and the cluster center:

$$d_t(y) = \|i_t - C_y\|, \quad (2)$$

where i_t represents the n input vector. C_y represents the y cluster center. $d_t(y)$ is the distance from i_t to C_y .

According to the minimum distance between the sample point and each cluster center, the data is divided into K pieces. The mean of the distance between all samples in each cluster is the new cluster center C'_y .

$$C'_y = \frac{1}{T_y} \sum_{P_y} I, \quad (3)$$

where P_y is the sample set of the y cluster center. T_y is the amount of data in the set. When the new clustering center is different from the initial clustering center, the above steps are repeated until the clustering center is fixed and the hidden layer node center C_y is obtained. The distance between node centers of each hidden layer is calculated and the minimum value is taken as the scale factor σ_y .

$$\sigma_y = \lambda \min_x \|C_y - qC_x\|. \quad (4)$$

In the formula, λ is the overlap coefficient, which is generally 1 initially and then adjusted through experiments.

Traditional k -means clustering algorithm requires K initial clustering centers. The minimum number of hidden layer nodes to reach the target precision is the number of hidden layer nodes of RBF neural network; then the center C_y and the width σ_y of hidden layer node are determined.

3. The Proposed Method

3.1. Linear Programming Determines Combination Weights. The standard model of general linear programming is shown in

$$\max k = \sum_{y=1}^t c_y i_y. \quad (5)$$

$$\text{s.t.} \begin{cases} \sum_{y=1}^t g_{xy} i_y = h_x & x = 1, 2, \dots, w \\ i_y \geq 0 & y = 1, 2, \dots, t \end{cases} \quad (6)$$

If solution $i = (i_1, i_2, \dots, i_t)$ satisfies (6), $i = (i_1, i_2, \dots, i_t)$ is said to be feasible. If $i = (i_1, i_2, \dots, i_t)$ simultaneously satisfies (5) to obtain the maximum value, it is called the optimal solution.

3.2. ARMA Model. Assume that the sequence $\{I_n, n = 0, \pm 1, \pm 2, \dots\}$ satisfies the following model:

$$I_n - \phi_1 I_{n-1} - \dots - \phi_u I_{n-u} = \varepsilon_n - \theta_1 \varepsilon_{n-1} - \dots - \theta_v \varepsilon_{n-v}. \quad (7)$$

I_n is zero mean stationary sequence. ε_n is stationary white noise with zero mean and variance σ_ε^2 . I_x is the ARMA

sequence, whose order is u and v . Applying the two operator polynomials $\phi(H)$ and $\theta(H)$, (7) can be written as $\phi(H)I_x = \theta(H)\varepsilon_n$.

For a general stationary sequence $\{I_n, n = 0, \pm 1, \pm 2, \dots\}$, set its average value $E(I_n) = \mu$. It satisfies the following model:

$$(I_n - \mu) - \phi_1(I_{n-1} - \mu) - \dots - \phi_u(I_{n-u} - \mu) = \varepsilon_n - \theta_1 \varepsilon_{n-1} - \dots - \theta_v \varepsilon_{n-v}. \quad (8)$$

If the operator polynomials $\phi(H)$ and $\theta(H)$ are used, (8) can be expressed as $\phi(H)(I_x - \mu) = \theta(H)\varepsilon_n$.

The steps for building the ARMA model are as follows.

3.2.1. AIC Criterion for ARMA Model Grading. Akaike information criterion (AIC) [21] determines the order of quasiside by selecting u and v , which satisfies the following equation:

$$\min AIC = t \ln \hat{\sigma}_s^2 + 2(u + v + 1). \quad (9)$$

In (9), t represents the capacity of the selected sample. $\hat{\sigma}_s^2$ represents the expected value of variance σ_s , and the stationarity of time series is related to u and v . If AIC has the minimum value when $u = \hat{u}$, $v = \hat{v}$, sequence ARMA(u, v) is considered to have the best stationarity.

When sequence ARMA(u, v) contains unknown mean value parameter μ , the model is $\phi(H)(I_n - \mu) = \theta(H)\varepsilon_n$. In this case, the number of unknown parameters is $z = u + v + 2$. AIC criterion is expressed where u and v are selected to satisfy

$$\min AIC = n \ln \hat{\sigma}_\varepsilon^2 + 2(u + v + 2). \quad (10)$$

In fact, the minimum value of (10) above is the same as that of (9) above. When the values of (u, v) are set as (\hat{u}, \hat{v}) , the values of (9) and (10) are the smallest.

3.2.2. χ^2 Test of ARMA Model. For sequence AR(u), assuming $\hat{\phi}_1, \hat{\phi}_2, \dots, \hat{\phi}_u$ is the unknown parameter estimate, then the residual $\hat{\varepsilon}_n = X_n - \hat{\phi}_1 X_{n-1} - \dots - \hat{\phi}_u X_{n-u}$, $n = 1, 2, \dots, t$ (assuming $I_0 = I_{-1} = \dots = I_{1-u} = 0$) is denoted by

$$\eta_z = \frac{\sum_{n=1}^{t-z} \hat{\varepsilon}_n \hat{\varepsilon}_{n+z}}{\sum_{n=1}^z \hat{\varepsilon}_n^2}, \quad z = 1, 2, \dots, L. \quad (11)$$

where L is the mantissa of the autocorrelation function of $\hat{\varepsilon}_n$, and Ljung's χ^2 test statistic is

$$\chi^2 = t(t+2) \sum_{z=1}^L \frac{\eta_z^2}{t-z}. \quad (12)$$

3.3. The Improved Genetic Algorithm. The traditional genetic algorithm (GA) has a strong global search ability for all possible solution sets, though the local search ability is weak. As a result, the convergence rate becomes slow and it is difficult to approach the global optimal solution. In addition, traditional genetic algorithms only judge the merits and demerits of solutions

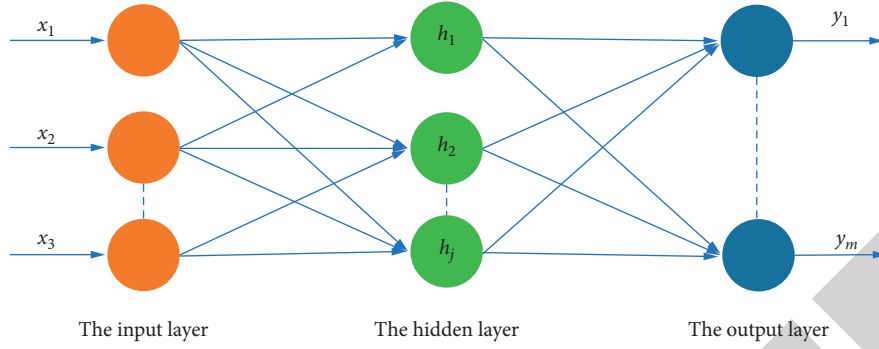


FIGURE 1: RBF neural network structure.

according to fitness. Consequently, in the early stage of evolution, some individuals' fitness value is too large, which leads to the population falling into the local optimal solution, resulting in premature convergence. Therefore, the crossover operator combined with genetic algorithm can change the global search ability, and the mutation operator can change the local search ability. Besides, the crossover probability and mutation probability of genetic operator can be adjusted adaptively according to the current individual fitness and the number of evolutionary iterations.

In traditional genetic algorithms, the crossover probability U_c is fixed. Generally, the maximum crossover probability U_{cmax} is not more than 0.8, and the minimum crossover probability U_{cmin} is not less than 0.3. If the selected individuals have a large crossover probability, the search scope of genetic algorithm will be expanded, and the global search ability will be enhanced. However, excessive crossover probability may lead to the destruction of the original high adaptability of chromosomes. If the crossover probability of selected individuals is small, the global search ability of genetic algorithm will be reduced and the convergence speed will be slow. Therefore, in the process of evolution, it is necessary to continuously adjust the crossover probability according to the current individual fitness and the number of evolutionary iterations.

In the early stages of evolution, individuals are usually less able to adapt to their environment. Therefore, when the individual fitness value is lower than the average fitness value, a larger crossover probability should be selected to increase the global search range of the algorithm. In the later stage of evolution, with the increasing number of iterations, individuals with low environmental adaptability have been gradually eliminated. Hence, when the individual fitness value is higher than the average fitness value, the crossover probability can be appropriately reduced according to the number of iterations, thus reducing the global search ability of the algorithm. Based on the above ideas, the improved adaptive crossover probability is

$$U'_c = \begin{cases} U_{cmax}, & F_{max} < F_{mean} \\ U_{cmax} - \frac{U_{cmax} - U_{cmin}}{iter_{max}} \times iter, & F_{max} \geq F_{mean} \end{cases}, \quad (13)$$

where $iter$ and $iter_{max}$ represent the current iteration number and maximum iteration number of the algorithm,

respectively, F_{max} represents the maximum fitness value of the two individuals to be crossed in the parent population, and F_{mean} represents the average fitness value of all individuals in the parent population. The fitness F was determined using the distorted form of the error function E between the actual output and the target output of the IGA-RBF network. The fitness of the F individual can be expressed as

$$F(i) = \frac{1}{E(i)}. \quad (14)$$

In traditional genetic algorithm, mutation probability U_w is fixed. Generally, the maximum variation probability U_{wmax} is not more than 0.1, and the minimum variation probability U_{wmin} is not less than 0.001. If the mutation probability is small, some important genes of the good chromosomes produced by crossover operation can be retained. However, if the mutation probability is large, the algorithm will always be in the state of random search, making it difficult for the excellent individuals produced by crossover operation to continue genetic operation. Therefore, in the process of evolution, it is necessary to continuously adjust the mutation probability according to the current individual fitness and the number of evolutionary iterations.

Similar to the crossover process, in the early stage of evolution, when the fitness value of an individual is lower than the average fitness value, small mutation probability is selected to preserve the excellent genes in the chromosome as much as possible. In the later stage of evolution, when the individual fitness value is higher than the average fitness value, the mutation probability can be appropriately increased according to the number of iterations, so as to improve the local search ability of the algorithm. Therefore, the adaptive mutation probability is put forward as

$$U'_w = \begin{cases} U_{wmin}, & F < F_{mean} \\ U_{wmin} + \frac{U_{wmax} - U_{wmin}}{iter_{max}} \times iter, & F \geq F_{mean} \end{cases}, \quad (15)$$

where $iter$ and $iter_{max}$ represent the current iteration number and maximum iteration number of the algorithm, respectively, F represents the fitness value of chromosomes to be mutated in the parent population, and F_{mean} represents

the average fitness value of all chromosomes in the parent population.

Since it is not guaranteed that the new individuals, which will be generated after crossover or mutation, have better environmental adaptability than the parent individuals, it is necessary to compare the fitness of the parent individuals and offspring individuals. Among them, the paternal individual represents the chromosome before crossover or mutation, and the offspring individual represents the chromosome after crossover or mutation. If the fitness value of the offspring is greater than that of the parent, it will indicate that the fitness of the offspring increases after crossover or mutation, and the offspring can continue to be genetically operated. On the contrary, if the fitness of the individual decreases after crossover or mutation, the offspring will be discarded and the parent will be used to replace the offspring to continue the genetic operation.

3.4. The Optimized RBF Neural Network with IGA. It is, in the process of training and optimization of RBF neural network, mainly to learn and adjust the center of hidden layer activation function, expansion constant, and continuous weight from hidden layer to output layer. For solving these network parameters, the learning algorithm of the network is different because of the different determination methods. Therefore, there are various learning algorithms of RBF neural network, such as self-organizing center selection method and orthogonal least square method. According to the different methods of selecting the center of the activation function of the hidden layer, the learning algorithm of RBF neural network can be divided into two types. (1) Input samples are used as the center of the activation function, which will not change once it is determined, such as orthogonal least square method. (2) The activation function center is not fixed, but can be dynamically adjusted continuously through RBF neural network training, such as self-organizing center selection method. Since the self-organizing center selection method needs to determine the number of hidden layer nodes manually, the number of hidden layer nodes has a direct influence on the determination of activation function center and expansion constant. The unreasonable number of hidden layer nodes may lead to the deterioration of RBF neural network performance, nonlinear approximation ability, and generalization ability. Therefore, this paper chooses orthogonal least square method, which is widely used and can automatically design a minimum network structure, as the learning algorithm of RBF neural network.

In order to reduce the complexity of adjusting parameters during network training, IGA was used to optimize the orthogonalization sequence of RBF neural network based on orthogonal least square method. Orthogonal least square method orthogonalizes the column vectors of the output matrix of the hidden layer through Gram-Schmidt [22] orthogonalization, so as to find the column vectors that contribute the most energy to the output of the network. Although the difference in the order of orthogonalization will lead to the change of the column vectors that contribute

the most energy to the output, the order of selecting the hidden layer activation function centers from the input samples will also change. However, the total training error of RBF neural network does not change because of its different order. Therefore, the improved genetic algorithm can be used to optimize the column vector of output matrix of hidden layer of RBF neural network on the premise of satisfying the target error. According to the optimal column vector, the optimal orthogonalization sequence is determined. Finally, the optimal activation function center and hidden layer node number are obtained, and the IGA-RBF neural network with better structure is designed, so as to reduce the tedious degree of adjusting parameters in network training. Specific steps using IGA optimization based on orthogonal least squares are shown below.

Step 1: initialize the network. All input sample matrices $I = [i_1, i_2, \dots, i_t]$ were taken as the center of the activation function, and the number of nodes in the hidden layer was initialized to t . At the same time, set the maximum iteration number l_{max} of the genetic algorithm, and set the initial iteration number l to 0. When an algorithm iteration is completed, the value of l automatically increases by 1.

Step 2: compute the output matrix $\Phi = [\varphi_1, \varphi_2, \dots, \varphi_t]$ of the hidden layer, where $\varphi_x = \varphi(I, c_x)$ is the output vector of the x hidden layer element. When training sample i_z is input, the output of the x hidden layer unit is $\varphi(i_z, c_x)$, also known as the activation function of the hidden layer. It is usually expressed by Gaussian function, which is expressed by

$$\varphi(i_z, C_x) = \exp\left(\frac{-\|i - c_x\|^2}{2\sigma_x^2}\right), \quad (16)$$

where c_x is the center of the activation function and σ_x is the expansion constant.

Step 3: since the problem to be solved is to find the optimal order of t column vectors of Φ , by sorting t different column vectors, $t!$ possible solutions can be obtained to form the solution space of the problem in question. Genetic algorithm can not directly optimize the solution space data to solve the problem, so it needs to map it into the genetic space by means of gene coding. First, the different sequences of t column vectors of Φ are numbered successively, and then these numbers are processed by binary encoding so that they can represent the solution of the problem to be solved in the form of 0/1 strings.

Step 4: initialize the population. From the solution space of the problem to be solved, the possible solution sets of groups μ were randomly selected to form the initial parent population $U_0 = [u_1, u_2, \dots, u_\mu]$, where $u_x (x = 1, 2, \dots, \mu)$ is a set of possible solutions to the problem, as well as a set of orthogonalized sequences.

Step 5: the negative gradient steepest descent method is used to continuously modify the connection weight ω_{xy} between the hidden layer and the output layer. When

the total error E of the network is less than the target error ε , the correction is stopped.

Step 6: use the genetic operator.

Step 6.1: to calculate the individual fitness value, according to the individual fitness value and roulette selection criteria, two better individuals u_x and v_y were selected from the parent population.

Step 6.2: cross u_x and v_y according to the improved crossover probability U'_c to generate daughter chromosomes.

Step 6.3: mutate the daughter chromosomes generated in Step 6.2 according to the improved mutation probability U'_w .

Step 7: when the iteration number l of the algorithm exceeds the set maximum iteration number, the genetic evolution process is planted. In this case, you can determine the number of nodes at the hidden layer. Otherwise, proceed to Step 5 to continue the next generation of genetic operations.

Step 8: if the termination conditions of the genetic algorithm are met, the chromosome with higher fitness value can be selected as the most individual in the solution space through the calculation of fitness function. At this point, the center of the network activation function can be determined according to the most individual obtained, namely the optimal orthogonalization order. As the expression form of this individual is genotype string structure, decoding operation is required, that is, the inverse operation of binary coding.

3.5. Linear Optimization Method to Determine the Weight.

The key steps of AR-RBF combined prediction model are as follows.

- (1) Obtain the required data, given the training data set of AR-RBF model, including the sample set of input data and output data.
- (2) The feature values are obtained by feature extraction, standardization, and normalization of the original data of shooting topics.
- (3) Calculate the absolute value of relative prediction error $re_{xy} = |(j_{xy} - n_y)/n_y|$ of each unit model prediction method (ARMA time series model and improved RBF neural network prediction model), where re_{xy} is the absolute value of error of variable y in model x . n_y is the expected output of the y variable. j_{xy} is the actual output value of the y variable in the x model. $x = 1, 2, \dots, w$, and $y = 1, 2, \dots, w$ are the number of prediction models, and T is the number of prediction model variables.
- (4) Linear optimal solution is used to determine the weighting coefficient of AR-RBF model. Such that the weighting coefficient i_x ($x = 1, 2, \dots, w$) satisfies $\min k = \sum_{y=1}^w |\sum_{x=1}^w i_x j_{xy} - n_y|$, where k is the total value of prediction error of the model.

- (5) After the weighted factor i_x is calculated, the weighted factor is used to recalculate the predicted value \hat{b} of the RBF combined prediction model $\hat{b} = \sum_{x=1}^w i_x j_{xy}$.

4. Experimental Results and Analysis

The experimental environment of this algorithm is Win10 system, with Intel dual-core processor, 8G memory, and 500G hard disk. The data were obtained from Weibo and Baidu, and the Terminal High Altitude Area Defense (THAAD) incident in South Korea was taken as a hot topic. Also, the respective Baidu index and Weibo index are obtained. The duration of the relationship was 1 month.

4.1. *The Effectiveness of the Proposed Algorithm.* The RBF neural network and the proposed algorithm are selected to compare their prediction accuracy, so as to verify the prediction effect of the proposed algorithm. Figure 2 shows the comparison of RBF neural network with the predicted click-through number calculated in this paper and the actual click-through number of Weibo. R square value, $RMSE$ value, and MSE value were used as evaluation indexes to measure goodness of fit between predicted value and actual value. The equation is shown in the following equation:

$$R^2 = 1 - \frac{\sum_x (\hat{j}_x - j_x)^2}{\sum_x (\hat{j}_x - j_x)^2}$$

$$RMSE = \sqrt{\frac{1}{w} \sum_{x=1}^w (j_x - \hat{j}_x)^2}, \quad (17)$$

$$MSE = \frac{1}{w} \sum_{x=1}^w (j_x - \hat{j}_x)^2,$$

where w is the total number of samples. j is the true value, and \hat{j} is the predicted value. $RMSE$ is the root mean square error and MSE is the mean absolute error.

The number of microblog clicks is a direct reflection of the changing barometer of sentiment. As can be seen from Figure 2, the evolution of the heat wave is characterized by rapid outbreak and slow decline, experiencing four stages of germination, acceleration, maturity, and decline, basically in line with the life cycle of online public opinion. To be specific, BP neural network in the stage of budding and the prediction results are not accurate and the delay is serious, especially in the outbreak of love and its predicted value deviates from the real value are large. As for the algorithm in this paper, the predicted trend not only is consistent with the real situation in a relatively flat place, but also can accurately reveal the trend of the real value when the wave peak is large, which shows, that based on the improved RBF neural network algorithm, we can mine the internal law of love communication from the increasing media data. At the same time, the development trend of the situation is predicted accurately, and the error of the predicted result is smaller

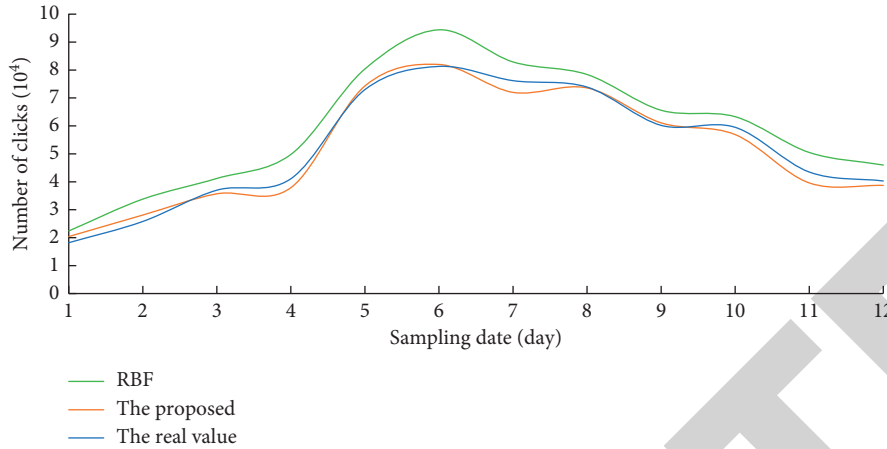


FIGURE 2: The predicted results with RBF neural network.

TABLE 1: Prediction results of the proposed algorithm and RBF neural network.

Methods	R^2	RMSE	MSE
RBF	0.8336	0.0286	0.00075
The proposed	0.9361	0.0148	0.00018

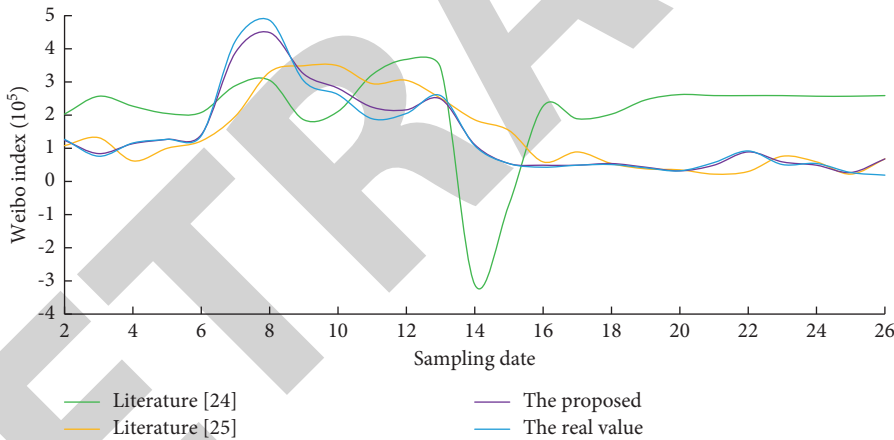


FIGURE 3: The predicted results of THAAD incident on Weibo index.

than the actual value. The R square value, RMSE value, and MSE value of the prediction results of BPF neural network algorithm and the algorithm in this paper are shown in Table 1.

Comparative analysis of the data in Table 1 shows that each index of the proposed algorithm is significantly better than that of RBF neural network algorithm. Its R square value is 0.9361, higher than the BRF neural network algorithm 0.8336. Compared with BRF neural network, the RMSE value of the proposed method is decreased by 48.2%. MSE value is greatly reduced compared with BRF neural network. Experimental results display that the proposed

algorithm has higher prediction accuracy and is an effective method for public opinion trend prediction.

4.2. *The Comparative Analysis with Other Algorithms.* Three algorithms are adopted to predict the Weibo index of the THAAD incident, as shown in Figure 3.

As can be seen from Figure 4, the fitting effect of literature [23] is the worst, and the prediction results obtained cannot be well fitted to public opinion data. There is a big gap between the sum effect of literature [24] and the trend of public opinion, and the trend effect of public opinion

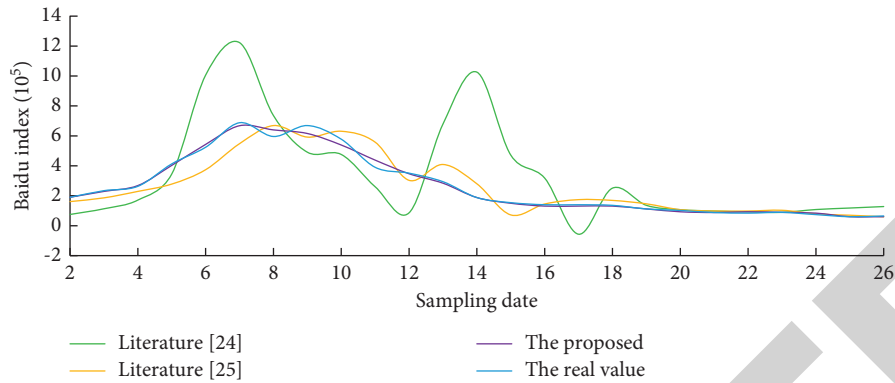


FIGURE 4: The predicted results of THAAD incident on Baidu index.

obtained is average. The residual error fitting effect of the algorithm in this paper is good and the trend of public opinion is close to the actual trend of public opinion.

Three algorithms are used to predict the Baidu index of the THAAD incident, as shown in Figure 4.

As can be seen from Figure 4, the fitting effect of literature [23] is poor, and it is unable to control the change rule of the trend of and situation well, and the predicted results are unstable. The prediction results obtained by literature [25] and the algorithm in this paper are better, but the prediction results obtained by literature [24] are not as good as the prediction results of the algorithm in this paper. The prediction results of the target prediction value of the algorithm in this paper are the most accurate among the three algorithms.

5. Conclusion

With the rapid development of mobile communication technology and Internet technology, social hot issues have been rapidly spread and entered the public view. Consequently, network public opinion data becomes increasingly complex, while detailed public opinion analysis and accurate prediction of public opinion development are of particular importance. At present, the prediction research of news public opinion is still in the development stage. In addition, there are still a great number of problems to be solved. This paper proposes an algorithm of news public opinion prediction based on IGA-RBF neural network. The algorithm uses Weibo index and Baidu index as news public opinion data sources to study the news public opinion trend of "THAAD incident." Firstly, the improved BRF neural network and ARMA time series are trained, and the weight of AR-RBF combined prediction model is determined by one linear optimal method. The experimental results indicate that the proposed algorithm not only has a good fitting effect, but also has a small deviation. The target prediction value of the proposed algorithm is more accurate, providing more reliable data for news public opinion prediction. In the future, the author will conduct prediction experiments on more news platforms and make a more systematic analysis of the performance of the algorithm in this paper.

Data Availability

The labeled dataset used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the Ministry of Education Humanities and Social Sciences Research Youth Foundation "Research on Improving the Public Opinion Guidance of Party Media Based on Big Data" (Project No. 17YJCZH023).

References

- [1] L. Liu, S. Yu, X. Wei, and Z. Ning, "An improved Apriori-based algorithm for friends recommendation in microblog," *International Journal of Communication Systems*, vol. 31, no. 2, p. e3453, 2018.
- [2] M. Lim, "Freedom to hate: social media, algorithmic enclaves, and the rise of tribal nationalism in Indonesia," *Critical Asian Studies*, vol. 49, no. 3, pp. 411–427, 2017.
- [3] S. Yang, *We-media as a Marketing Tool for Fashion Business in China: Case*, LingSiYaoYao Ltd, Chennai, Tamilnadu, 2020.
- [4] C. Zhou, F. Su, T. Pei et al., "COVID-19: challenges to GIS with big data," *Geography and sustainability*, vol. 1, no. 1, pp. 77–87, 2020.
- [5] B. O'Connor, R. Balasubramanian, B. R. Routledge, and N. A. Smith, "From Tweets to Polls: Linking Text Sentiment to Public Opinion Time series," in *Proceedings of the Fourth International AAAI Conference on Weblogs and Social media*, Washington, DC USA, May 2010.
- [6] A. Hermida, F. Fletcher, D. Korell, and D. Logan, "Share, like, recommend: decoding the social media news consumer[J]," *Journalism Studies*, vol. 13, no. 5-6, pp. 815–824, 2012.
- [7] Q. Jia, Y. Guo, G. Wang, and S. J. Barnes, "Big data analytics in the fight against major public health incidents (including COVID-19): a conceptual framework," *International Journal of Environmental Research and Public Health*, vol. 17, no. 17, p. 6161, 2020.
- [8] M. Kaigo, "Social media usage during disasters and social capital: twitter and the Great East Japan earthquake[J]," *Keio communication review*, vol. 34, no. 1, pp. 19–35, 2012.

- [9] L. Mora, R. K. R. Kummitha, and G. Esposito, "Not everything is as it seems: d," p. 101599, 2021.
- [10] Z. Cai, H. Huang, and D. Ma, "A System Dynamics Model for Public Opinion Diffusion and Response Strategy of Government Creditability Crisis Event during the COVID-19 Pandemic," in *Proceedings of the 2021 the 5th International Conference on Management Engineering*, pp. 124–131, Software Engineering and Service Sciences, Chengdu, China, January 2021.
- [11] A. Tandon, A. Dhir, I. Almgren, G. N. AlNemer, and M. Mäntymäki, *Fear of Missing Out (FoMO) Among Social media Users: A Systematic Literature Review, Synthesis and Framework for Future research*, Internet Research, Bingley, UK, 2021.
- [12] X. Chen, X. Lan, and J. Wan, P. Lu and M. Yang, "Evolutionary prediction of nonstationary event popularity dynamics of weibo social network using time-series characteristics," *Discrete Dynamics in Nature and Society*, vol. 2021, Article ID 5551718, 19 pages, 2021.
- [13] P. Jiang, B. Wang, H. Li, and H. Lu, "Modeling for chaotic time series based on linear and nonlinear framework: application to wind speed forecasting," *Energy*, vol. 173, pp. 468–482, 2019.
- [14] X. Liu and C. Liu, "Information diffusion and opinion leader mathematical modeling based on microblog," *IEEE Access*, vol. 6, pp. 34736–34745, 2018.
- [15] A. Miglani and N. Kumar, "Deep learning models for traffic flow prediction in autonomous vehicles: a review, solutions, and challenges," *Vehicular Communications*, vol. 20, p. 100184, 2019.
- [16] Q. Zhang, J. Gao, J. T. Wu, Z. Cao, and D. Dajun Zeng, "Data science approaches to confronting the COVID-19 pandemic: a narrative review," *Philosophical Transactions of the Royal Society A*, vol. 380, no. 2214, p. 20210126, 2021.
- [17] H. Bai and G. Yu, "A Weibo-based approach to disaster informatics: incidents monitor in post-disaster situation via Weibo text negative sentiment analysis," *Natural Hazards*, vol. 83, no. 2, pp. 1177–1196, 2016.
- [18] F. Jia and C. C. Chen, "Emotional characteristics and time series analysis of Internet public opinion participants based on emotional feature words," *International Journal of Advanced Robotic Systems*, vol. 17, no. 1, p. 1729881420904213, 2020.
- [19] S. Shayaa, N. I. Jaafar, S. Bahri et al., "Sentiment analysis of big data: methods, applications, and open challenges," *IEEE Access*, vol. 6, pp. 37807–37827, 2018.
- [20] Z. Majdisova and V. Skala, "Radial basis function approximations: comparison and applications," *Applied Mathematical Modelling*, vol. 51, pp. 727–743, 2017.
- [21] J. E. Cavanaugh and A. A. Neath, "The Akaike information criterion: background, derivation, properties, application, interpretation, and refinements," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 11, no. 3, p. e1460, 2019.
- [22] P. Shakya, M. S. Kulkarni, and A. K. Darpe, "Bearing diagnosis based on Mahalanobis-Taguchi-Gram-Schmidt method," *Journal of Sound and Vibration*, vol. 337, pp. 342–362, 2015.
- [23] X. G. Chen, S. Duan, and L.-d. Wang, "Research on trend prediction and evaluation of network public opinion," *Concurrency and Computation: Practice and Experience*, vol. 29, no. 24, p. e4212, 2017.
- [24] X. Chen, M. Xia, J. Cheng, X. Tang, and J. Zhang, "Trend Prediction of Internet Public Opinion Based on Collaborative filtering," in *Proceedings of the 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, pp. 583–588, IEEE, Changsha, China, August, 2016.
- [25] J. V. Roberts and L. J. Stalans, *Public Opinion, Crime, and Criminal justice*, Routledge, England, UK, 2018.