Egyptian Informatics Journal 24 (2023) 149–160

Contents lists available at ScienceDirect

Egyptian Informatics Journal

journal homepage: www.sciencedirect.com

Prediction of complex public opinion evolution based on improved multi-objective grey wolf optimizer

Yilin Su [a], Yongsheng Li [b,*], Shibin Xuan [c]

[a] School of Electronic Information, Guangxi Minzu University, Nanning 530006, Guangxi, China
[b] School of Artificial Intelligence, Guangxi Minzu University, Nanning 530006, Guangxi, China
[c] School of Artificial Intelligence, Guangxi Minzu University, Nanning 530006, Guangxi, China

ARTICLE INFO

Article history:
Received 29 December 2022
Revised 31 January 2023
Accepted 12 February 2023
Available online 18 February 2023

Keywords:
Gray wolf optimizer
Multiobjective
Complex public opinion
Lotka-Volterra model

ABSTRACT

In an emergency, predicting the evolution trends of network public opinion is of great significance for governments to manage the situation and associated social public opinion trends. Most of the existing public opinion prediction models have the disadvantages of long time consumption, limited types of predicted public opinion categories and low-accuracy prediction results. To improve the ability to predict the evolutionary trend of complex public opinion, an improved multiobjective gray wolf optimizer (IMOGWO) is proposed. IMOGWO uses logistic and Lotka-Volterra models to initialize the wolf population and improve the population validity; it designs a nonlinear function to adjust the population update factor to improve the exploration and exploitation ability and the local search ability of the wolf population; and it introduces the elite retention policy and a Pareto-optimal solution set to achieve multiple objectives using the principle of a nondominated solution set. In simulation experiments, various network emergencies were taken as empirical analysis cases. The experimental results show that the improved multiobjective GWO has good accuracy and universality on complex public opinion evolution prediction and can better predict the evolution trends of various types of complex public opinion than other tested methods.

1. Introduction

Network public opinion refers to prominent community public opinion with specific views on social problems on the internet, which is a construct of the expression of social public opinion. On the web, opinions about the occurrence, improvement and exchange of certain hot topics and focal point issues, which are generally expressed with the enthusiastic aid of the public, have an impact on opinions and inclinations via, for example, BBSs and Blogs. Due to the improvement of the internet, public opinion on the web is having an increased impact on governmental activities and options. After major emergencies, public opinion on the internet can rapidly develop into a powerful social force. Therefore, the effective guidance of online public opinion is not only related to the success or failure of on-site emergency response, but also to the stability and harmony of society [1].

Research on the prediction of evolutionary trends of public opinion with data from the internet is very much applicable the Chinese context, and many scholars have studied the evolutionary process of public opinion diffusion [2]. Research methods usually involve the use of mathematical statistical models, Backpropogation algorithm, traditional system theory methods, machine learning, etc. [3]. Yang Wenyang used a genetic algorithm to generate an optimization model for the public opinion spreading process to better plan the spreading path of online public opinion [4]. Nie Lisheng used kernel principal component analysis (KPCA) and the particle swarm random forest algorithm to predict public opinion trends, and this algorithm significantly improved the prediction accuracy compared with back propagation algorithms [5]. The paper [6] used a fused particle swarm and k-means clustering algorithm for microblog public opinion trend analysis and detection. Lv Z [7] constructed a model based on a neural network to solve the

* Corresponding author.
E-mail address: lyshlh@163.com (Y. Li).

Peer review under responsibility of Faculty of Computers and Information, Cairo University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.eij.2023.02.001
1110-8665/

defect of low prediction accuracy of traditional SVM and other models. Chen S et al. [8] compensated for the lack of a single prediction algorithm for public opinion trends and improved the accuracy of prediction. An algorithm based on an IGA-RBF neural network was established to address complex public opinion prediction based on the news. He J et al. [9] proposed the "digital twin of public opinion" model to realize the early warning of negative public opinion events.

The evolutionary development of public opinion is the result of mutual checks and balances of multiple factors, which exhibit complex nonlinear characteristics and spiraling trends [10], and the time-series data of public opinion often present time-series correlations and have a large data volume. By combining the relevant literature published at home and abroad in recent years, it is found that the research on complex public opinion prediction by domestic and foreign scholars is still in the exploratory stage, and there is a lack of research on complex public opinion prediction at home and abroad. Most of the existing research on public opinion prediction relies on time series analysis or neural networks. Complex public opinion prediction involves not only the prediction of the prevalence of a single public opinion but also the prediction of multi-information public opinion. Among the relevant studies on the prediction of the evolution of online public opinion, few researchers have studied public opinion events in which both public opinion information and intervention information about the information exist simultaneously. To address the shortcomings of the above studies, to solve the problem of predicting the evolutionary trend of public opinion and to improve the prediction speed and accuracy of the evolutionary trend of public opinion, this paper proposes an improved multiobjective gray wolf optimizer by combining various strategies. The gray wolf optimizer (GWO) has strong convergence performance, and there exists a convergence factor that can be adaptively adjusted as well as an information feedback mechanism, which has good performance in terms of the accuracy when solving the problem. The convergent GWO has good performance in terms of both accuracy and convergence speed, and it has higher performance than the particle swarm algorithm in exploring unknown and challenging search space problems. In this paper, we use logistic and Lotka-Volterra models to initialize the wolf population, improve GWO with multiobjective and added sine functions and logarithmic function optimization to fit the curves when public opinion changes suddenly, expand the search range and increase the convergence speed of the algorithm, and then verify the usability and stability of the algorithm by fitting experiments to several sudden events.

2. Principle introduction

2.1. Model of the evolution of complicated public opinion data

The public opinion statistics found in new media can be divided into two types: in the first, there is only one piece of public opinion in the same time interval, and in the second, multiple portions of public opinion over a brief time period.

There is solely-one piece of public opinion data for identical match at the same time interval.

For single network public opinion on an emergency, the evolution of public opinion is similar to the survival law of species, and population growth is composed of increasing, rapid growth, slow growth and extinction periods. The evolution of public opinion involves four stages: incubation period, outbreak period, steadiness length and recession duration [11].

The same event has multiple public opinion messages that simultaneously appear over a short time period.

In the event of an emergency, various types of widely disseminated information can be classified as public opinion information or intervention information, with the release of intervention information often lagging behind that of public opinion information. Considering the delayed release of intervention information, the communication of public opinion and intervention information is divided into two stages: the independent communication stage of public opinion information, and the joint communication stage of public opinion and intervention information.

Phase I: Public opinion information is spread separately. At the early stage of the emergency, only public opinion information exists in the public space. Increases in the number of supporters is similar to those found in the number of patients during a single type of mass attack in ecology, and propagation conforms to the logistic model.

Phase II: Public opinion and intervention information is co-disseminated. Given the influence of human-bounded rationality and the attractiveness of intervention information, any increase in the number of supporters of public opinion records is constrained by the number of people available and the aggressive impact of intervention information. Phase II's dissemination rules conform to the Lotka-Volterra competition model.

The propagation of complex public opinion follows a specific development pattern, including butterfly effects and collective reactions. Some public opinion events with correlation and overlapping audiences begin to merge with each other as online public opinion evolves continuously, so there does not exist a model with specific parameters that can predict the evolution direction of all public opinion trends; as an alternative, a mathematical model that conforms to the law of public opinion evolution can be applied, and then appropriate parameters are set according to the characteristics of each public opinion event. According to the laws and characteristics of complex public opinion evolution, to solve the difficulties caused by different types of public opinion arising from the same public opinion event before and after the turning point in the prediction of public opinion evolution and the multiple types of information about complex public opinion, this paper chooses to use the logistic model and Lotka-Volterra competition model to initialize the parameters of the predicted public opinion evolution model, iteratively update the prediction using the improved MOGWO parameters of the public opinion evolution model, and finally find the optimal model parameters.

2.2. Logistic model of a single population

The logistic curve model is a common "s" curve was first proposed by biologist Weihurst in his research on population growth. Weihurst found that population growth has bearing capacity as the initial population growth rate increases with time. After a period of rapid growth, the growth rate starts to gradually slow down after reaching a certain limit. Finally, the total social population attains a stable state [12]. The style of information at the stage of unbiased propagation of public opinion facts is comparable to the system described by the logistic curve model, so the improved logistic model was chosen to predict the evolution of public opinion trends at the separate dissemination stage of emergencies, as shown in formula (1):

$$\begin{cases} \frac{dx_1}{dt} = r_1 x_1 \left(1 - \frac{x_1}{K_1}\right) \\ x_1(t)|_{t=0} = x_1(0)\#\# \end{cases} \tag{1}$$

where $x_1(\mathrm{t})$ is the number of supporters won by public opinion information at time t, r_1 is the growth rate of the number of supporters of public opinion information, K_1 is the upper limit of the population of supporters won by public opinion information, and $x_1(0)$ is the initial number of supporters of public opinion information gathering.

Obtain formula (2) according to formula (1):

$$x_1(t) = \frac{K_1 x_1(0) e^{r_1 t}}{K_1 + x_1(0)(e^{r_1 t} - 1)} = \frac{K_1}{1 + \left(\frac{K_1}{x_1(0)} - 1\right) e^{r_1 t}} = \frac{K_1}{1 + a e^{-r_1 t}} \tag{2}$$

among $a = \frac{K_1}{x_1(0)} - 1$.

2.3. LV competition model

The Lotka-Volterra competition model [13] was developed by A Loka and Volterra, an Italian mathematician, who proposed the model based on the mutual feeding relationship between populations in a food chain. This model establishes a theoretical foundation for the competition between populations and is widely applied in ecology, systems science and informatics. The logistic model is not as suitable as the Lotka-Volterra competition model for describing the change rules of multiple populations.

After intervention information is released, growth in the population of supporters of public opinion information is inhibited by the competitive effect of the intervention information. Similarly, the growth of the supporters of intervention information is affected by their own growth potential, the potential for growth in the environment and competition for public opinion information. Therefore, we make use of the increased Lotka-Volterra model to predict the trends of online public opinion at the stage of public opinion and intervention data co-transmission. The mathematical model is:

$$\begin{cases} \frac{dx_1}{dt} = r_1 x_1 \left(1 - \frac{x_1}{K_1} - \alpha_1 \frac{x_2}{K_2}\right) \\ \frac{dx_2}{dt} = r_2 x_2 \left(1 - \frac{x_2}{K_2} - \alpha_2 \frac{x_1}{K_1}\right) \\ x_1(t)|_{t=t_0} = x_1(t_0) x_2(t)|_{t=t_0} = x_2(0) \end{cases} \tag{3}$$

x_2(t) indicates the number of supporters of public opinion information at time t, r_2 is the increase price of the range of supporters of public opinion information, and K_1 is the higher bound of the wide variety of supporters that public opinion facts can win; α_1, α_2 indicates that the inhibition of x_1 on x_2 is the attraction of intervention information at time t, and the inhibition of x_2 on x_1 is the attraction of public opinion information at time t; t_0 refers to the intervention information release time; x_1(t) is the population of supporters at public opinion information time t_0, and x_2(t) is the quantity of supporters at intervention information time t_0.

The degree of public awareness and its influence on public opinion and intervention information changes over time. With the passage of time, the attractiveness of public opinion information may gradually decline and tend toward zero. In contrast, the attractiveness of intervention information initially decreases and gradually increases to maximum attractiveness. The attractiveness of public opinion and intervention information is defined as follows:

$$\alpha_1 = \sigma(t) = \mu e^{-[\varepsilon(t-t_0)]^\beta} \tag{4}$$

$$\alpha_2 = \gamma(t) = \frac{\omega}{1 + e^{-[\rho + \varepsilon(t-t_0)]}} \tag{5}$$

where μ is the propagation intensity of public opinion information at moment t_0, β is an adjustment parameter such as public bounded rationality, ε is the intensity of intervention information release, ω is the upper limit of the attractiveness of intervention information, and ρ is the ability of intervention information at t_0 to win supporters at all times.

From (4)(5)(6), we can obtain:

$$\begin{cases} \frac{dx_1}{dt} = r_1 x_1 \left(1 - \frac{x_1}{K_1} - \frac{\omega}{1+e^{-[\rho+\varepsilon(t-t_0)]}} \frac{x_2}{K_2}\right) (6.1) \\ \frac{dx_2}{dt} = r_2 x_2 \left(1 - \frac{x_2}{K_2} - \mu e^{-[\varepsilon(t-t_0)]^\beta} \frac{x_1}{K_1}\right) (6.2) \\ x_1(t)|_{t=t_0} = x_1(t_0) x_2(t)|_{t=t_0} = x_2(0) (6.3) \end{cases} \tag{6}$$

$$\begin{cases} \frac{dx_1}{dt} = r_1 x_1 \left(1 - \frac{x_1}{K_1} - \frac{\omega}{1+e^{-[\rho+\varepsilon(t-t_0)]}} \frac{x_2}{K_2}\right) (6.1) \\ \frac{dx_2}{dt} = r_2 x_2 \left(1 - \frac{x_2}{K_2} - \mu e^{-[\varepsilon(t-t_0)]^\beta} \frac{x_1}{K_1}\right) (6.2) \\ x_1(t)|_{t=t_0} = x_1(t_0) x_2(t)|_{t=t_0} = x_2(0) (6.3) \end{cases} \tag{7}$$

2.4. Gray wolf optimizer

The gray wolf optimizer (GWO) is a new swarm intelligence algorithm that utilizes the potential of Miljalili et al. in 2014. It is flexible, scalable, and achieves accurate stability between exploration and development. The GWO simulates the conduct of a predatory gray wolf population in accordance with the cooperation mechanism among wolves. Each wolf performs a function in the population conforming to natural law and strict social systems [14].

In a GWO, wolves are sorted in accordance with their fitness. We can consider a wolf in the wolf pack as a viable answer and identify the wolves corresponding to the modern-day finest solution, superior answer and suboptimal answer as the α,β, and γ Wolf, respectively; the wolf with the lowest health ω is considered the lowest ranked individual. A wolf follows the motion of the most fulfilling answer and hunts in accordance with issued instructions.

During the search process, the predatory conduct of the gray wolf populace can be described as follows:

$$\mathrm{D} = \left|\mathrm{C} * \mathrm{X_p(t)} - \mathrm{X(t)}\right| \tag{7}$$

$$\mathrm{X(t+1)} = \mathrm{X_p(t)} - \mathrm{A} * \mathrm{D} \tag{8}$$

In formula (7), D is the distance between the wolf and the target. Formula (8) is the coordinate transformation of a wolf, X_P (t) is the position of the *t*-generation target, X (t) is the position of a single wolf in the *t*-generation wolf pack, A and C are coefficients, and the calculation formula is as follows:

$$a = 2 - 2 * \frac{iter}{Max_{iter}} \tag{9}$$

$$\mathrm{A} = 2\mathrm{a} * \mathrm{r_1} \tag{10}$$

$$\mathrm{C} = 2\mathrm{r_1} \tag{11}$$

where *iter* is the population of iterations,Max_{iter} is the largest quantity of iterations, and $r_1, r_2 \in [0, 1]$. When the wolf catches quarry (i.e., when it attacks its target), the replacement of the male or female gray wolf is decided in accordance with the three types of wolves, and the model is as follows:

$$\mathrm{D_i^j(t)} = \left|\mathrm{C} * \mathrm{X_i^j(t)} - \mathrm{X^j(t)}\right| \tag{12}$$

$$\mathrm{X_i^j(t+1)} = \mathrm{X_i^j(t)} - \mathrm{A} * \mathrm{D_i^j(t)} \tag{13}$$

$$\mathrm{X(t+1)} = \frac{1}{3} * \sum \mathrm{X_m(t+1)} \tag{14}$$

where $D_i^j(t)$ is the gap between the *t*-generation and $i(i = \alpha, \beta, \gamma)$ wolves. Formulas (12) - (14) define the ω wolf according to α,β,γ step length and the direction of wolf movement, respectively. Formula (14) represents the new era of gray wolves established after a location update.

3. Improved gray wolf optimizer

3.1. Composite model method initializes the population of the GWO

The evolution of public opinion information is divided into separate dissemination stages and common dissemination stages, and the public opinion in the two stages has different evolutionary characteristics. The overall public opinion data cannot be expressed by a linear equation, and it presents a nonlinear characteristic. Based on observations and research, we found that the change trend of data in the separate dissemination stage of public opinion information is very similar to the process portrayed by the logistic curve model, and the curve of the Lotka-Volterra model is similar to the change principle of online public opinion in the copropagation stage of public opinion information and intervention information, so to comprehensively address the characteristic factors of evolutionary dissemination of public opinion events, a combined logistic and Lotka-Volterra model is used to initialize the population of the improved gray wolf optimizer for predicting complex public opinion.

- Independent dissemination of public opinion information stage.

Changes in the trend of public opinion data at this stage is very similar to the population growth process described by the logistic curve model, so the logistic model is used to initialize the population of the GWO. At time 0-t_0, public opinion information conforms to the logistic competition model. We use the three sum method to calculate the initial value of model parameters $K_1, a,$ andr_1, divide the training set data into three equal groups and sum the reciprocal $x_1(t)$ of each group:

$$\begin{cases} S_1 = \sum_{t=1}^{R} \frac{1}{x_1(t)} = \frac{R}{K_1} + \frac{ae^{-r_1}\left(1-e^{-r_1 R}\right)}{K_1(1-e^{-r_1})} \\ S_2 = \sum_{t=R+1}^{R} \frac{1}{x_1(t)} = \frac{R}{K_1} + \frac{ae^{-(R+1)r_1}\left(1-e^{-r_1 R}\right)}{K_1(1-e^{-r_1})} \\ S_3 = \sum_{t=2R+1}^{R} \frac{1}{x_1(t)} = \frac{R}{K_1} + \frac{ae^{-(2R+1)r_1}\left(1-e^{-r_1 R}\right)}{K_1(1-e^{-r_1})} \end{cases} \tag{16}$$

According to formula(15), $K_1, a,$ andr_1 of formula (8), (9) and (10) can be obtained by:

$$r_1 = \frac{\ln\frac{S_1-S_2}{S_2-S_3}}{R} \tag{16}$$

$$K_1 = \frac{R}{S_1 - \frac{(S_1-S_2)^2}{(S_1-S_2)-(S_2-S_3)}} \tag{17}$$

$$a = \frac{(S_1 - S_2)^2(1 - e^{-r_1})K_1}{[(S_1 - S_2) - (S_2 - S_3)]e^{-r_1}(1 - e^{-Rr_1})} \tag{18}$$

$$\mathrm{X}(0) = [K_1, a, r_1] \tag{19}$$

- Public opinion and intervention information simultaneously existing and spreading together.

The changing of public opinion and intervention information data at this stage conform to the Lotka-Volterra model, so the Lotka-Volterra model is used to initialize the population of the GWO.

Public opinion information conforms to the Lotka-Volterra competition model from moment t_0.

① Obtaining r_1 and K_1 after substitution in formula(6.1), the gray wolf modeling method is used to transform formula(6.1) into:

$$\frac{dx_1(t)}{dt} = x_1(t+1) - x_1(t) \tag{20}$$

$$x_1(t+1) - x_1(t) = r_1\frac{x_1(t+1)+x_1(t)}{2} - \frac{r_1}{K_1}\left[\frac{x_1(t+1)+x_1(t)}{2}\right]^2 - \frac{\frac{r_1\omega}{4K_2}[x_1(t+1)+x_1(t)][x_2(t+1)+x_2(t)]}{1+e^{-[\rho+\varepsilon(t-t_0)]}} \tag{21}$$

$$\frac{x_1(t+1) - x_1(t) - r_1x_1(t+1) + \frac{r_1x_1^2(t+1)}{K_1}}{x_2(t+1)} = \frac{\frac{\omega}{K_2}}{1+e^{\left(\varepsilon t_{0-\rho}\right)}e^{-\varepsilon t}} \tag{22}$$

② The three sum method is used to estimate $\frac{\omega}{K_2}e(\varepsilon t0 - 9)$and $\varepsilon$for (21).

③ The parameters obtained by ② are substituted into (20), and the least squares method is used to re-estimate $K_1r_1 and \frac{\omega}{K_2}$:

$$\left[r_1, \frac{r_1}{K_1}, \frac{\omega}{K_2}\right] = \left(B_1^TB_1^{-1}\right)^{-1}B_1^TX_1 \tag{23}$$

$$\begin{bmatrix} \frac{x_1(t_0+2)+x_1(t_0+1)}{2} & -\left[\frac{x_1(t_0+2)+x_1(t_0+1)}{2}\right]^2 & -\frac{[x_1(t_0+2)+x_1(t_0+1)][x_2(2)+x_2(1)]}{4} \\ \frac{x_1(t_0+3)+x_1(t_0+2)}{2} & -\left[\frac{x_1(t_0+3)+x_1(t_0+2)}{2}\right]^2 & -\frac{[x_1(t_0+3)+x_1(t_0+2)][x_2(3)+x_2(2)]}{4} \\ \vdots & \vdots & \vdots \\ \frac{x_1(t_0+n)+x_1(t_0+n-1)}{2} & -\left[\frac{x_1(nt_0+)+x_1(t_0+n-1)}{2}\right]^2 & -\frac{[x_1(t_0+n)+x_1(t_0+n-1)][x_2(n)+x_2(n-1)]}{4} \end{bmatrix} \tag{24}$$

$$X_1 = [x_1(t_0+2) - x_1(t_0+1) \;\; x_1(t_0+3) + x_1(t_0+2) \;\; \cdots \;\; x_1(t_0+n) - x_1(t_0+n-1)]^T \tag{25}$$

④ The expression $x_1(t+1)$ is obtained according to the discrete form of the Lotka-Volterra model:

$$x_1(t+1) = \frac{e^{r_1}x_1(t)}{1 + \frac{(e^{r_1}-1)x_1(t)}{K_1} + \frac{\omega(e^{r_1}-1)x_2(t)}{r_1K_2}} \tag{26}$$

From moment t_0, intervention information conforms to the Lotka-Volterra competition model.

① The gray wolf modeling method uses a differential equation to describe the dynamic behavior of the research object and directly estimates the parameters in the equation, so it is suitable for parameter estimation in the Lotka-Volterra competition model. Substituting the above $r_1$and K_1 into formula(6.2), we use the gray wolf modeling method to convert formula (6.2) into:

$$\frac{dx_2(t)}{dt} = x_2(t+1) - x_2(t) \tag{27}$$

$$x_2(t+1) - x_2(t) = r_2\frac{x_2(t+1)+x_2(t)}{2} - \frac{r_2}{K_2}\left[\frac{x_2(t+1)+x_2(t)}{2}\right]^2 - \frac{\mu[x_1(t+1)+x_1(t)][x_2(t+1)+x_2(t)]}{4K_1e^{[\varepsilon(t-t_0)]^\beta}} \tag{28}$$

② The least squares method is often used for curve (polynomial) fitting, so it can be used to solve the above parameters. We substitute ε and K_1 obtained from the public opinion model into formula (28), set $\beta = 1$, and use the least squares method(formula (29)) to estimate $K_2, r_2,$ and μ.

$$\left[r_2, \frac{r_2}{K_2}, \mu\right] = \left(B_2^T B_2^{-1}\right)^{-1} B_2^T X_1 \tag{29}$$

$$B = \begin{bmatrix} \frac{x_2(2)+x_2(1)}{2} & -\left[\frac{x_2(2)+x_2(1)}{2}\right]^2 & -\frac{[x_1(2)+x_1(1)][x_2(2)+x_2(1)]}{4K_1 e^{2\varepsilon}} \\ \frac{x_2(3)+x_2(2)}{2} & -\left[\frac{x_2(3)+x_2(2)}{2}\right]^2 & -\frac{[x_1(3)+x_1(2)][x_2(3)+x_2(2)]}{4K_1 e^{3\varepsilon}} \\ \vdots & \vdots & \vdots \\ \frac{x_2(n)+x_2(n-1)}{2} & -\left[\frac{x_2(n)+x_2(n-1)}{2}\right]^2 & -\frac{[x_1(t_0+n)+x_1(t_0+n-1)][x_2(n)+x_2(n-1)]}{4K_1 e^{n\varepsilon}} \end{bmatrix} \tag{30}$$

$$X_2 = \left[x_2(2) - x_2(1) \quad x_2(3) - x_2(2) \quad \cdots \quad x_2(n) - x_2(n-1)\ \right]^T \tag{31}$$

The initial population parameters of two wolves are obtained:

$$X_1(t_0) = [K_1, r_1, \omega, \varepsilon] \tag{32}$$

$$X_2(t_0) = [K_2, r_2, \beta, \mu] \tag{33}$$

where $X_1(t_0)$ is the initial population parameter of the public opinion information, and $X_2(t_0)$ is the initial population parameter of the intervention information. The proposed algorithm improves upon the initialization population method to address the overly random and scattered initialization time points of the original GWO so that the initial wolf colony is within a feasible range, which effectively reduces optimization time and increases the possibility of finding the optimal solution (Fig. 1).

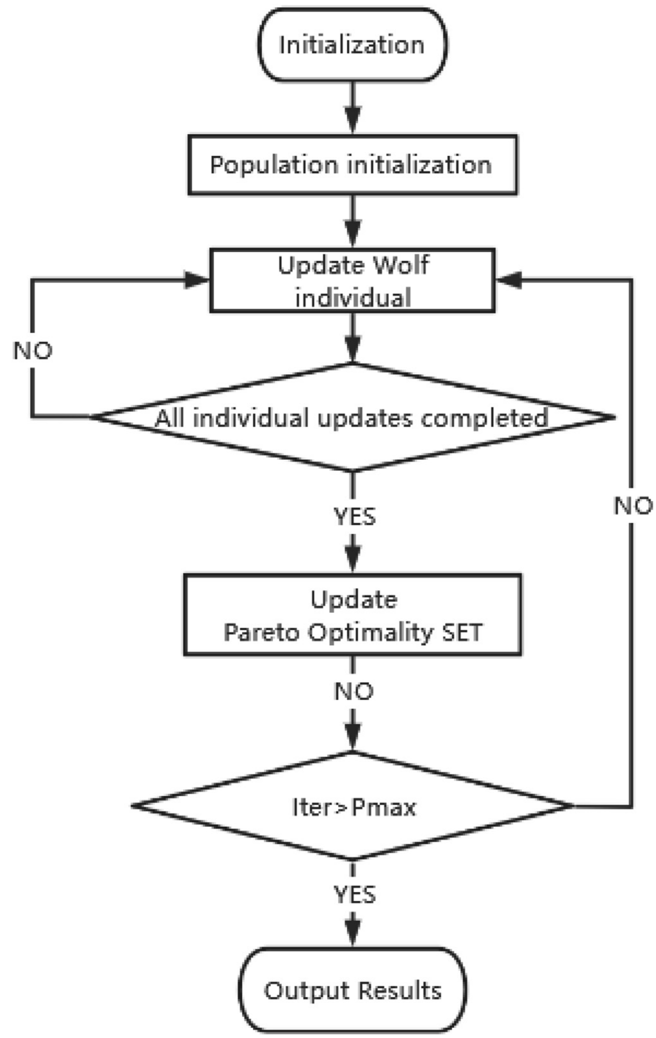

Fig. 1. Experimental steps.

3.2. Nonlinear function to adjust the population update factor and improve the search efficiency

The convergence factor is ɑ, and step size in the gray wolf optimizer is A_i. As two important variation factors in the population update formula, they directly affect the performance of GWO. Modifying the transformation of these two factors to nonlinear convergence variation makes the population update of GWO with an improved convergence factor and step length factor have the characteristics of nonlinear variation; hence, it becomes a more suitable population update method for complex opinion prediction problems and can help the algorithm converge faster and find the optimal solution with the highest accuracy. The strategy of linearly decreasing convergence factor ɑ and single random step size in the original algorithm cannot reflect the actual optimization process, so a non-linear change strategy is used for the convergence factor ɑ and step size A_i. In this paper, trigonometric and logarithmic functions are introduced on the basis of the original algorithm, and the improvement ideas of the grey wolf algorithm are as follows:

To coordinate the wolf population through international and neighborhood search, it is necessary to adjust convergence angle a. The population diversity is excellent at the initial stage of optimization, so it should quickly decrease at the early stage of optimization and slowly decrease at the later stage of optimization. We balance the exploration and mining capacity of the algorithm and combine it with the population of iterations $(\frac{iter}{Max\_iter})$. When the wide variety of instances increases, a regularly decreases, and the search location progressively concentrates around the location where the best answer is most likely to occur to enhance search efficiency. The function log() has fast exchange in the initial stage and gradual and mild exchange in the later stage. Therefore, ln() based completely on formula(9) is used to make the convergence element exhibit a steep nonlinear convergence exchange at the initial stage and mild exchange at the late stage. The multiplied F (34) is generated as follows:

$$a = 2 - 2 * \ln\left(1 + \left(\frac{iter}{Max_{iter}}\right)^{\frac{1}{2}}\right) \tag{34}$$

Utilizing the mathematical bounds of the sinusoidal function, we stabilize the capacity of the optimizer for international exploration and nearby improvement via object searching, increase the fluctuation range, and use nonlinear modifications in regular variation to add the sine feature based completely on formula(10) to achieve formula(36). To accommodate the enhancement of sin(), the fixed coefficient of formula(11) is modified to π extended with the aid of irregular numbers to obtain formula(35), so that its search vicinity is modified into a irregularly banded area, and search effectiveness is improved. Based on formula(12) - (14) and together with formula(34) - (36), step exchange components are updated to formula (37) and (38):

$$C_i = (2r_{1,i} - 1) * \pi \tag{35}$$

$$A_i = r_{2,i} * \sin(C_i) + A_i \tag{36}$$

$$D_i = |A_i * X_i(t) - X(t)| \tag{37}$$

$$X_i(t+1) = X_i(t) + a * D_i \tag{38}$$

where π is pi,$r_{1,i}, r_{2,i} \in [0,1], i = \alpha, \beta, \gamma$.

3.3. Elite retention policy and Pareto-optimal solution set to achieve multiple objectives

Multi-information opinion prediction in complex opinion forecasting is a multiobjective problem, and the main objective is to maximize the fit of multiple opinion event models with different model parameters that are interrelated in the same period, so this paper optimizes the original single-objective GWO into a multiobjective GWO.

Multiobjective optimization refers to the need to address multiple objectives simultaneously. The greatest advantage of multiobjective optimization over single objective optimization is the ability to deal with multiple conflicting objectives simultaneously and to obtain a set of Pareto-optimal solutions. In a multiobjective problem, a Pareto-optimal solution [15] is defined as follows: for any one solution in the solution set, no other solution can be found in the variable space that is superior to that solution; i.e., if all objective function values are superior, then the solution is the Pareto-optimal solution. The multiobjective optimization problem contains many subobjectives, and each subobjective has the possibility of conflicting with each other, and the metrics between subobjectives may also be different. It is impossible to obtain the optimal solution for each objective function under a certain decision vector.

In this paper, to convert the single-objective gray wolf optimizer algorithm into a multiobjective gray wolf optimizer, the elite retention strategy of NSGA-II is introduced [16–17]. The parent population is combined with the offspring individuals generated by global search and local search operations to form a large new population, and the new population is ranked using the nondominated ranking method and crowding degree to select the top N (N = population size) best gray wolf individuals as the next generation population. In addition, the hierarchical nondominated solution ranking method reduces the computational complexity of the algorithm, and the calculation of the crowding degree can preserve the solutions with lower similarity to maintain the diversity of the solution space and make the distribution of individuals on the current Pareto front as wide and uniform as possible, which helps to maintain the best individuals and improve the overall evolution of the population.

4. Experimental simulations

We carried out simulation experiments to evaluate the overall performance of the proposed algorithm. The experimental results show that the proposed algorithm has robust potential and can quickly discover the best solution. The running environment of the experiments consisted of: Windows 7 64 bit OS; software version: MATLAB R2018a; CPU: Inter (R) Core (TM) i5-5200U; Processor frequency: 2.20 GHz; RAM: 4 GB. To verify the efficiency of the proposed algorithm, simulation experiments were performed in different parameter simulation environments.

4.1. Experimental steps

① Initialize the parameters of the public opinion and intervention models, i.e., initialize the gray wolf position, and calculate Fit (t), i.e., the error.

② Find the three wolves α, β, γ and calculate convergence factor a.

③ Obtain the optimal solution set according to the Pareto optimal solution principle and assume that the three optimal solutions represent the α, β and γ wolf.

④ Subject the GWO to iterative optimization and repeat steps ②-④ until the maximum number of iterations of a single generation is reached.

⑤ Calculate whether the error of the current estimated fitting model is within the specified error range. If not, restart from ②.

4.2. Benchmarked on benchmark functions

To verify the effectiveness of the proposed IMOGWO algorithm, as shown in Table 1, this section uses 10 benchmark functions from different types of standard test functions used to evaluate the optimization problem for testing, where Dim denotes the dimensionality of the function, Range is the boundary of the function search space, and f_{min} is the most efficient value. The selected functions mainly include single-peaked, multipeaked, and fixed-dimensional multipeaked functions, as shown in Table 1. F1-F4 are single peak functions, the single peak function has a unique global optimum, which can be used to evaluate the exploitation capability of IMOGWO. F5-F10 are multipeaked benchmark functions and fixed dimensional multipeaked functions, the multipeaked function has more than one local extremum, which can be used to evaluate the balance between exploitation and exploration. The fixed-dimensional multipeaked function has more extreme value points, but due to the lower dimensionality, exploration is easier; therefore, this function can be used to test the stability of the algorithm. IMOGWO is run 30 times on each benchmark function, and the best outcomes (average value (ACE) and standard deviation (STD)) are statistically obtained, as shown in Table 2. To validate the results, IMOGWO was compared with PSO [18] and GWO.

Table 1
Benchmark functions.

Function	Dim	Range	f_{min}
$F_1(x) = \sum_{i=1}^{n} x_i^2$	30	[100,100]	0
$F_2(x) = \sum_{i=1}^{n} \lvert x_i\rvert + \prod_{i=1}^{n} \lvert x_i\rvert$	30	[-10,10]	0
$F_3(x) = \sum_{i=1}^{n} (\sum_{j-1}^{i} x_j)^2$	30	[-100,100]	0
$F_4(x) = max_i\{\lvert x_i\rvert, 1 \le i \le n\}$	30	[-100,100]	0
$F_5(x) = -20\exp\left(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}\right) - \exp(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)) + 20 + e$	30	[-32,32]	0
$F_6(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos(\frac{x_i}{\sqrt{i}}) + 1$	30	[-600,600]	0
$F_7(x) = \sum_{i=1}^{11}\left[a_i - \frac{x_1(b_i^2+b_i x_2)}{b_i^2+b_i x_3+x_4}\right]^2$	4	[-5,5]	0.0003
$F_8(x) = -\sum_{i=1}^{4} c_i \exp(-\sum_{i=1}^{6} a_{ij}(x_j - p_{ij})^2)$	3	[1,3]	−3.86
$F_9(x) = -\sum_{i=1}^{5}\left[(X - a_i)((X - a_i))^T + c_i\right]^{-1}$	4	[0,10]	−10.1532
$F_{10}(x) = (\frac{1}{500} + \sum_{j=1}^{25}\frac{1}{j+\sum_{i=1}^{2}(x_i - a_{ij})^6})^{-1}$	2	[-65,65]	1

Table 2
Results of benchmark functions.

F	STD			ACE		
	IMOGWO	GWO	PSO	IMOGWO	GWO	PSO
F1	4.2696e-44	1.9166e-29	7.2924e-05	2.5946e-41	1.2407e-27	7.2924e-05
F2	1.0962e-26	1.6745e-17	0.076348	2.0475e-25	9.7629e-17	0.076348
F3	4.944e-13	7.0033e-08	96.8513	5.3421e-07	2.5457e-05	96.8513
F4	1.9406e-11	4.92e-08	1.3644	1.2356e-09	7.0107e-07	1.3644
F5	7.9936e-15	7.9048e-14	0.0079421	1.4981e-14	1.0688e-13	0.0079421
F6	0	0	4.7564e-06	0.0069865	0.0075117	4.7564e-06
F7	0.00031052	0.00030884	0.00085138	0.0038292	0.0076924	0.00085138
F8	−3.321	−3.322	−3.322	−3.2734	−3.2353	−3.2031
F9	−10.153	−10.1531	−2.6829	−9.8131	−9.6462	−2.6829
F10	0.998	0.998	6.9033	3.844	5.235	6.9033

Fig. 2–11 depict the best statistical results obtained by running IMOGWO, GWO and PSO on the benchmark functions F1-F10 30 times. These outcomes confirm the performance of IMOGWO in solving various benchmark functions and prove the effectiveness of IMOGWO.

According to the results in Table 2, IMOGWO is able to provide very competitive results. It outperforms the other algorithms on the single peak functions F1-F4. The single peak functions are suitable for benchmark test development, so these results show that IMOGWO has superior performance in terms of developing optimal conditions. In addition, IMOGWO is able to provide very competitive results on F5-F10.The algorithm outperforms PSO and GWO on multimodal functions, and IMOGWO shows very competitive results. These results show the value of studying IMOGWO.

4.3. Single public opinion information experiment

In this chapter, the algorithm of this paper is compared with six mainstream algorithms, the algorithm in [11], the algorithm in [19], the algorithm in [20], PSO_RBF, standard RBF, and standard BP, in fitting five different real-event predictions. Each algorithm is run 500 times independently, and their relative errors obtained in each real event are recorded for each run, as shown in Figs. 12–16. The statistical results (relative errors) are shown in Table 3.

Figs. 12-16 show the experimental results of applying the algorithm in this paper and the six comparison algorithms to five different real events. It is easy to observe from the figures that the algorithm of this paper basically fits the data trend of 5 different real events, and the 500 independently ran experiments converge to a concentrated and small interval. This verifies the ability of the algorithm to converge toward the real data trend of real events, and the algorithm of this paper achieves these results in all experiments. Moreover, the algorithm converges within 5 s in almost every experiment, which proves that IMOGWO in this paper is effective. It is proven that IMOGWO in this paper is effective in improving the convergence accuracy and convergence speed of the algorithm, which suggests that the algorithm in this paper has a strong fitting prediction ability on single-opinion information events.

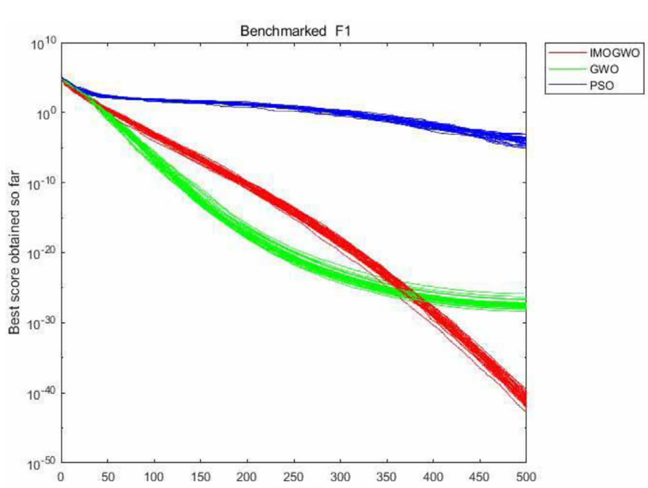

Fig. 2. BST F1.

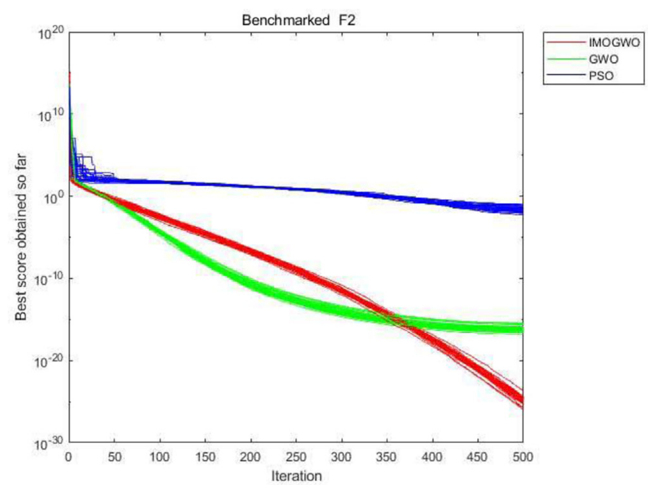

Fig. 3. BST F2.

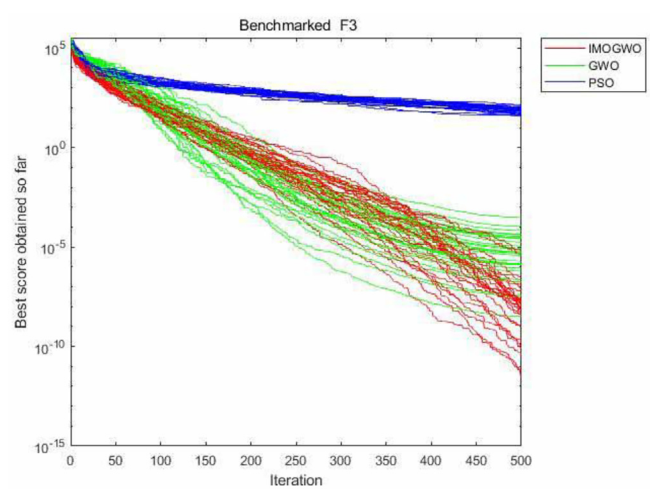

Fig. 4. BST F3.

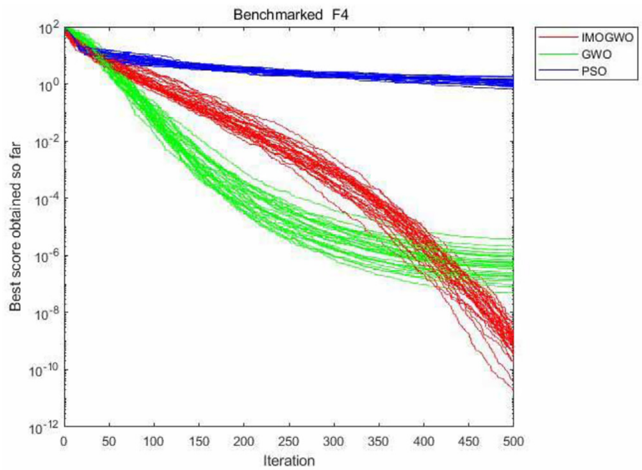

Fig. 5. BST F4.

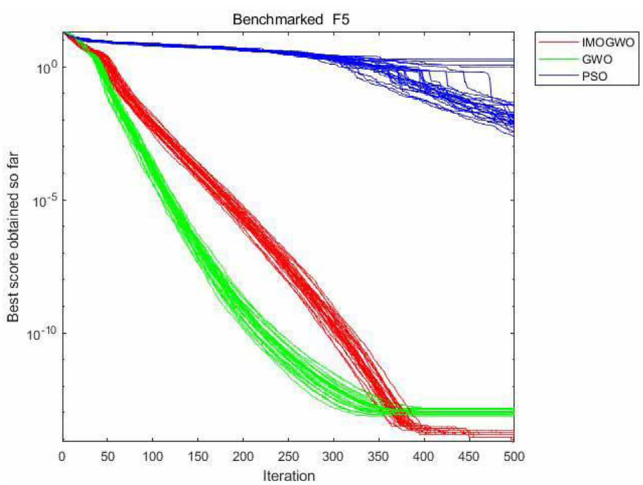

Fig. 6. BST F5.

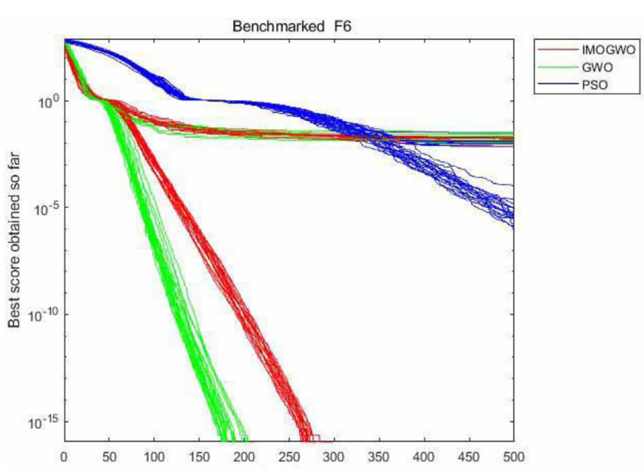

Fig. 7. BST F6.

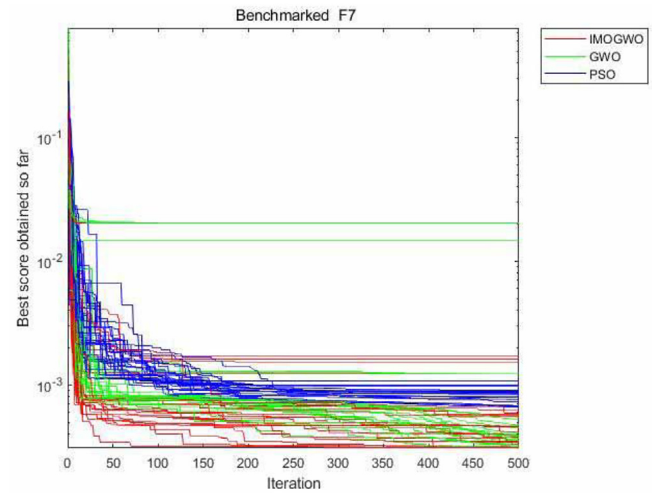

Fig. 8. BST F7.

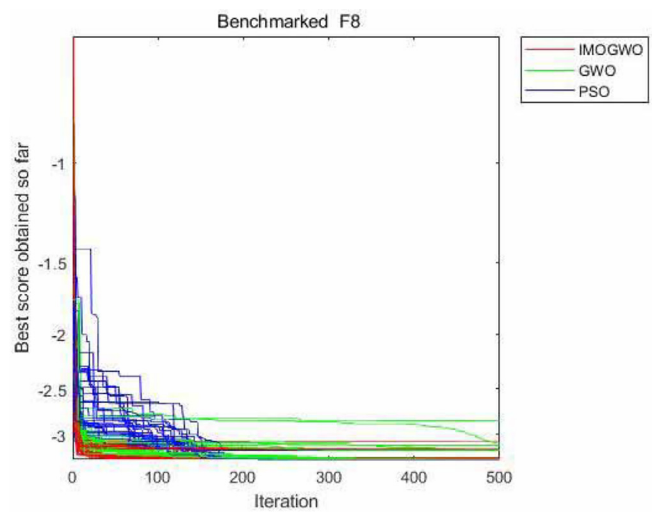

Fig. 9. BST F8.

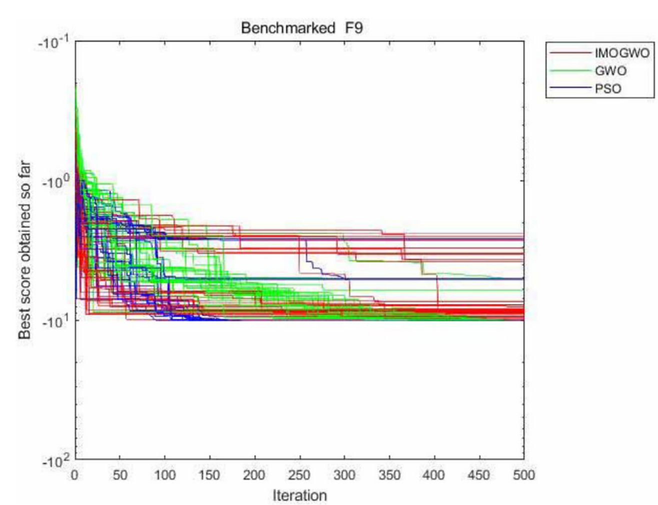

Fig. 10. BST F9.

Table 3 shows the relative error results of the algorithm in this paper and other algorithms in five experiments on fitting real events with single-opinion information. The relative error reflects the credibility of the measurement results, and the smaller the relative error is, the higher the credibility of the experimental results.

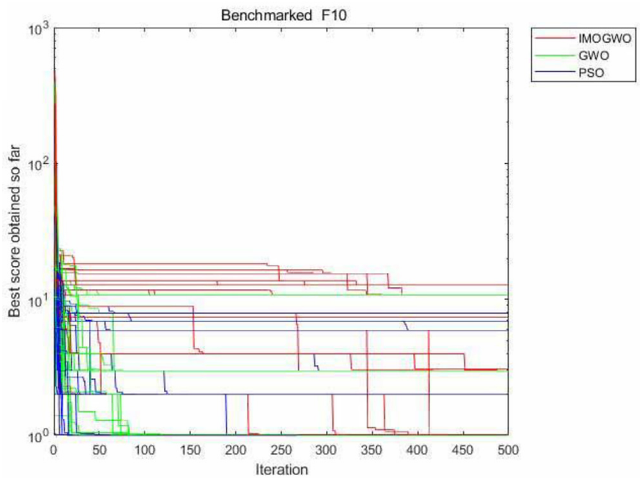

Fig. 11. BST F10.

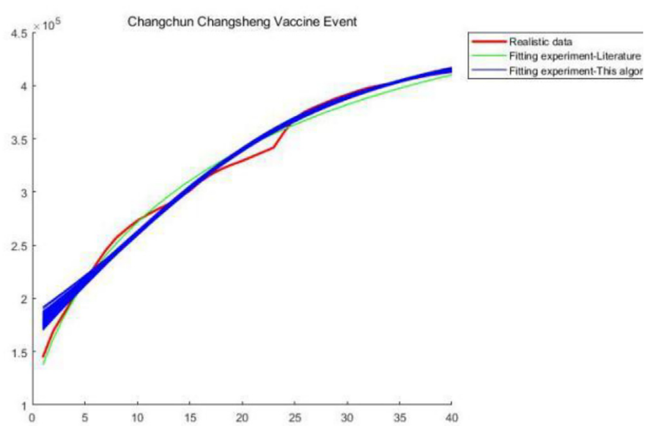

Fig. 12. Changchun Changsheng Vaccine Event.

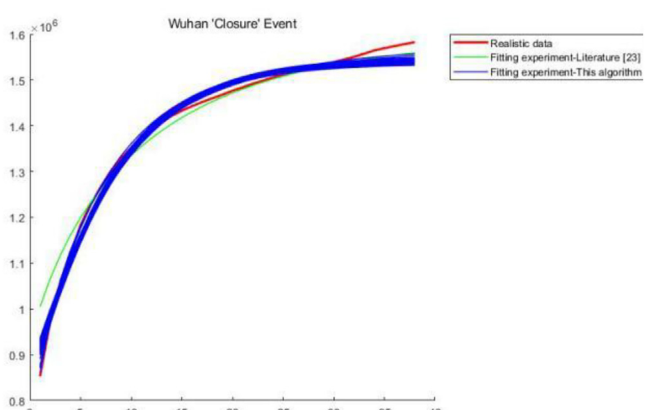

Fig. 13. Wuhan "Closure" Event.

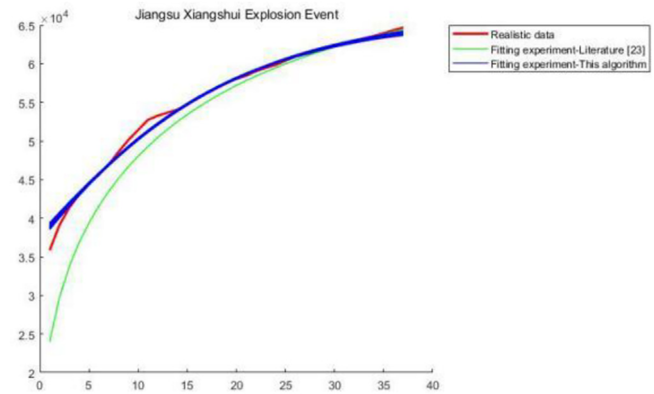

Fig. 14. Jiangsu Xiangshui Explosion Event.

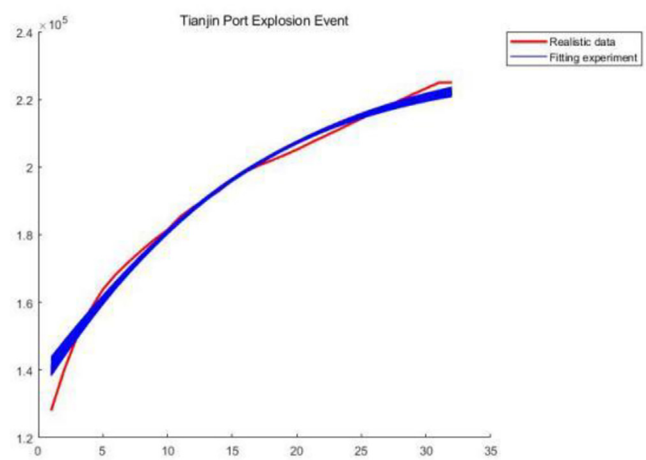

Fig. 15. Tianjin Port Explosion Event.

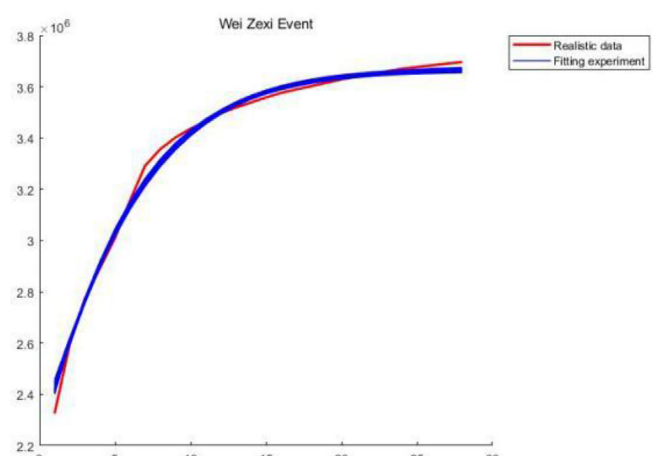

Fig. 16. Wei Zexi Event.

From the relative errors obtained after applying the seven algorithms to the five events in Table 3, we can see that the relative errors obtained by the algorithms in this paper in fitting the trends of the five real events are all smaller than the relative errors obtained by the algorithms in [11,18], and [19] and PSO_RBF for the same events, which indicates that the algorithms in this paper have a superior prediction fitting ability to events with single-opinion information.

4.4. Multiple information source public opinion information experiment

As shown in Table 4, seven parameters of the four-group opinion model K_1, r_1, x_0, t_0, μ, ρ and ε, and seven parameters of the intervention information model K_2, r_2, y_0, t_0, ω, β and ε, were set to generate four different sets of multi-information opinion event data. Then, the statistical results (mean relative error (MRE), mean

Table 3
Relative error of single public opinion information curve fitting experiments.

	Literature [11]	Proposed algorithm	Literature [19]	Literature [20]	PSO_RBF	Standard RBF	Standard BP
Changchun Changsheng Vaccine Event	0.0196	0.0175	/	/	/	/	/
Wuhan “Closure” Event	0.0164	0.0127	/	/	/	/	/
Jiangsu Xiangshui Explosion Event	0.0340	0.0056	/	/	/	/	/
Tianjin Port Explosion Event	/	0.0105	0.0820	/	/	/	/
Wei Zexi Event	/	0.0068	/	0.0774	0.1026	0.1782	0.1222

Table 4
Parameters of the Experimental Model.

	Public opinion model parameters							Intervention model parameters						
parm	K_1	r_1	x_0	t_0	μ	ρ	ε	K_2	r_2	y_0	t_0	ω	β	ε
1	50,000	2	2	2	1.2	0.6	0.6	50,000	2	2	2	0.4	0.5	0.6
2	40,000	3	4	4	1.5	0.3	0.2	40,000	5	6	4	0.7	0.6	0.2
3	44,500	4.31	2	4	0.151	1.533	2.003	573,130	2.28	5	4	1.702	1.014	2.003
4	34,200	3.31	3	4	1.556	0.333	1.263	34,200	5.26	4	4	0.212	0.654	1.263

absolute percentage error(MAPE), coefficient of determination(R^2)) obtained by fitting 100 predictions for each group of multi-information opinion events using the modified MGWO are shown in Table 5. MRE is a statistical indicator used to measure the error between the predicted and actual values, MAPE is a statistical indicator used to measure the accuracy of the prediction, and R^2 is used to determine the degree of approximation of the model to the sample data.

Figs. 17, 19, 21 and 23 show the results of this paper's algorithm in fitting the trends of the public opinion models from Experiment 1 to Experiment 4 in Table 4, respectively, and Figs. 18, 20, 22 and 24 show the results of this paper's algorithm in fitting the trends of the intervention models from Experiment 1 to Experiment 4 in Table 4, respectively. It can be seen that this paper's algorithm can obtain an almost perfect fitting for both the public opinion model and the intervention model, and it can accurately fit the model directions and model data, which indicates that this paper's algorithm is superior in predicting public opinion events with multiple information.

Table 5 lists the MRE, MAPE, and R^2 values obtained by the algorithm of this paper in fitting and predicting four groups of multi-

Table 5
Fitting error for the multiple information source public opinion information experiment.

	Public opinion model			Intervention model		
	MRE	MAPE(%)	R^2	MRE	MAPE(%)	R^2
1	0.0086	0.1379	1.0002	0.0189	0.2647	1.0001
2	0.0181	0.2902	0.9997	0.0097	0.1162	1.0005
3	0.1002	1.6039	0.9992	0.73441	0.0088	1
4	0.0026	0.0413	1	0.0019	0.0229	1.0002

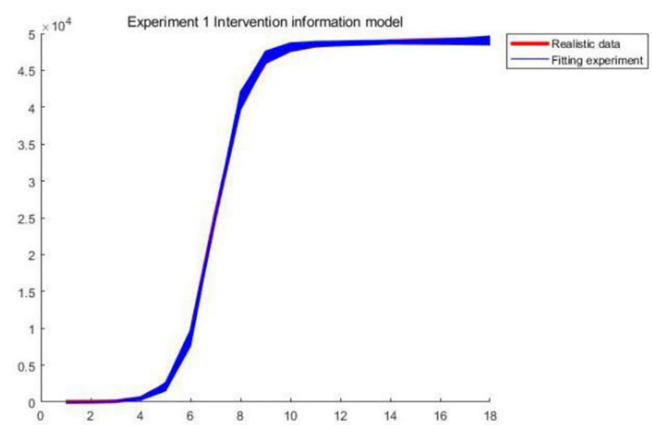

Fig. 18. Intervention Model in Exp1.

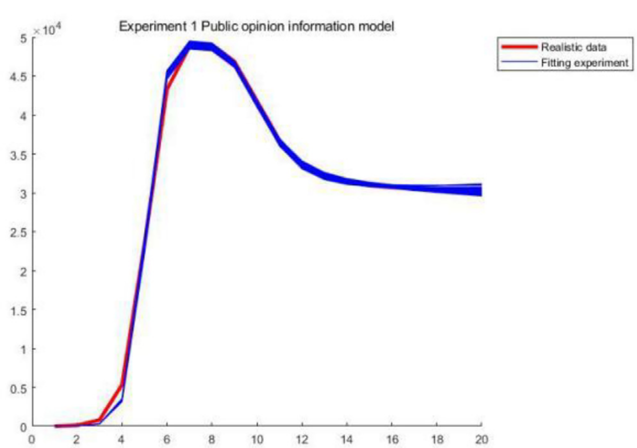

Fig. 17. Public opinion model in Exp1.

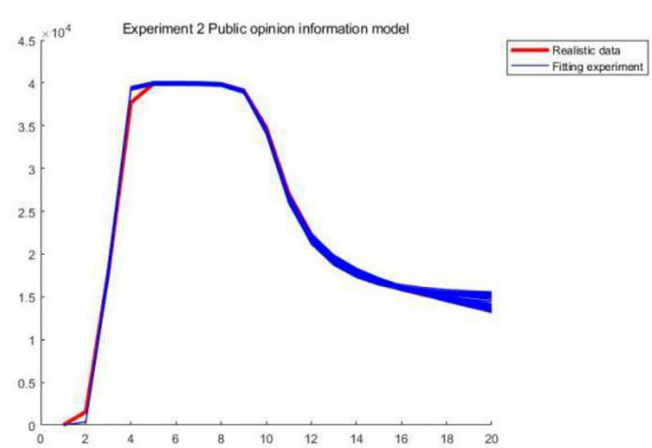

Fig. 19. Public Opinion Model in Exp2.

information opinion data 100 times, where the smaller the MRE and MAPE are, the better the experimental effect of the algorithm is, and the closer R^2 is to 1, the better the fit obtained by the algorithm. From the table, it can be seen that the algorithm of this paper has good performance in four groups of multi-information public opinion data, and the MRE and MAPE for the four groups of multi-information public opinion data are basically less than 0.1, the MAPE is basically less than 1%, and R^2 is very close to 1. This indicates that the algorithm of this paper has high accuracy and strong stability in fitting public opinion events and has an excellent fitting ability.

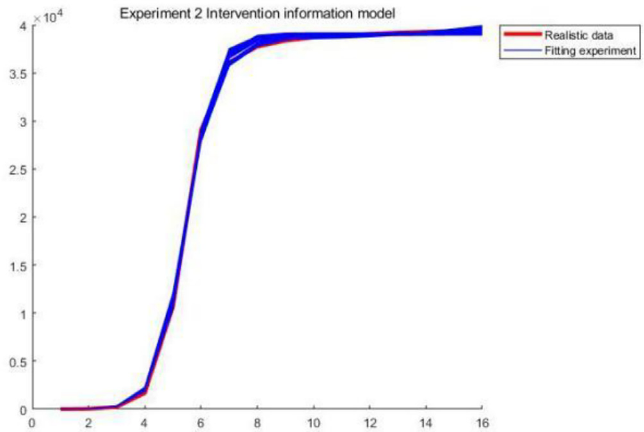

Fig. 20. Intervention Model in Exp2.

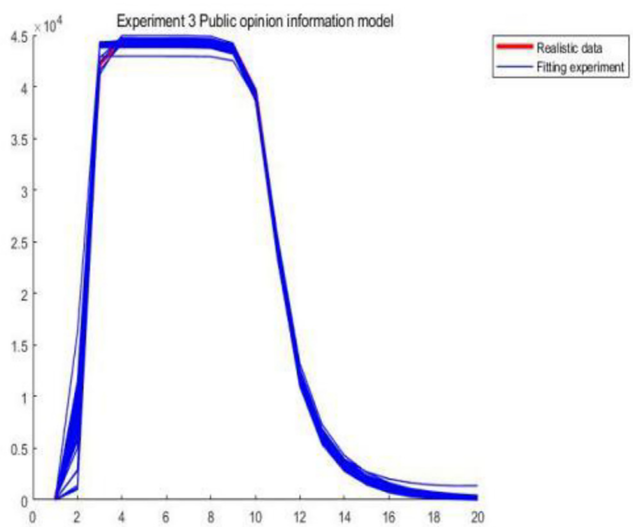

Fig. 21. Public Opinion Model in Exp3.

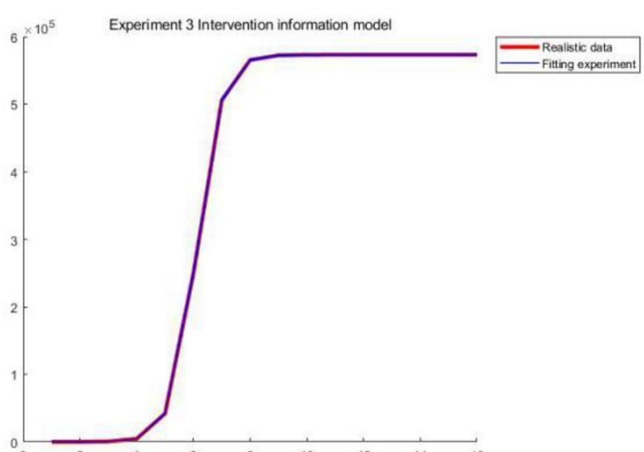

Fig. 22. Intervention model in Exp3.

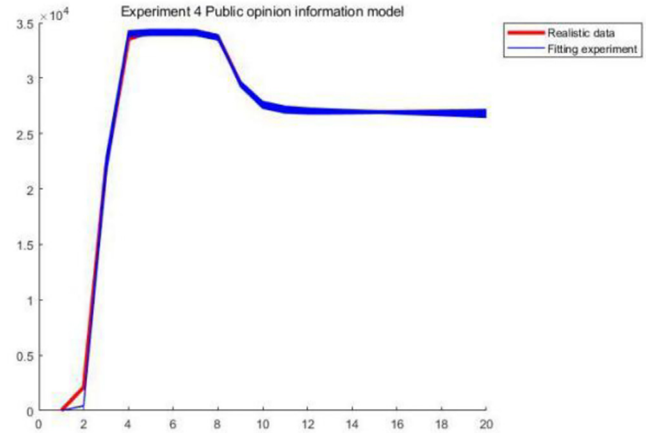

Fig. 23. Public Opinion Model in Exp4.

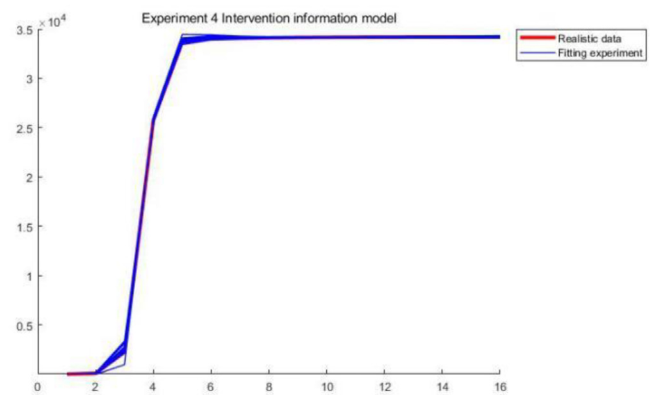

Fig. 24. Intervention model of Exp4.

5. Conclusion

Internet public opinion events are characterized by rapid spread, wide coverage and the involvement of a large number of people. If improperly handled, they can induce illegal and extreme behaviors that threaten social stability. Especially in the context of the current COVID-19 pandemic, this poses increasingly new challenges to the management of online public opinion. The proposed algorithm has low time cost and good optimization of the public opinion model of unequal competition, i.e., the public opinion intervention information competition model. The multiple fitting simulation results conducted with complex parameters show that the proposed algorithm outperforms the other curve fitting algorithms tested. In the future, we will consider the practical and rapid discovery of ideal model parameters in a dynamic and complex environment with heightened throughput and accuracy. The proposed strategy improves model fitting, reducing the time and cost required while maintaining strong application value.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

This work was supported by the National Natural Science Foundation of China (No. 61866003).

References

[1] Tie-jun C, Cong N, Lan-ping F. The evaluation of internet public opinion guidance effect for serious emergencies based on confidence level. Inform Sci 2022;40(02):118–26.
[2] Wang X, Wu J, Liu C, et al. Fault time series prediction based on lstm recurrent neural network. J Beijing Univ Aeronaut Astronaut 2018;44(4):13.
[3] Zhitao W, Yixue X. Situational awareness analysis of network public opinion based on evolutionary modeling. J Intelligence 2022;41(09):71–8.
[4] Wen-yang Yang. Research on the construction of optimization model of network public opinion propagation based on genetic algorithm and greedy algorithm. J China Acad Electron Inform Technol 2020;15(11):1057–64.
[5] Lisheng NIE. Research on trend prediction of public opinion based on KPCA and particle swarm random forest algorithm. Modern Electron Techn 2019;42(15):79–82.
[6] Jin Hu. Application of fusion PSO algorithm and K-means algorithm in microblogging public opinion monitoring. Hunan Univ 2018. doi: https://doi.org/10.27135/d.cnki.ghudu.2018.000040.
[7] Lv Z, Jan MA. Prediction of the forwarding volume of campus microblog public opinion emergencies using neural network. Mob Inf Syst 2022;2022:1–8.
[8] Chen S, Li G, Chen C-L. Research on prediction of news public opinion guiding power based on neural network. Secur Commun Networks 2022;2022:1–9.
[9] He J, Qi Y, Feng J, Xiang A, Liu Z. Towards public opinion digital twin: A Conceptual prototype. Sci Program 2022;2022:1–13.
[10] Luo G, Zhang Z, Bao Z, et al. A novel method for the time series data processing and analysis of social networks. J Comput 2020;31(4):273–87.
[11] Tie-jun Cheng, Man Wang. Network public opinion trend prediction of emergencies based on variable weight combination. Comput Sci 2021;48(S1). pp. 190–195+202.
[12] Qiu X, Zhao L, Wang J, Wang X, Wang Q. Effects of time-dependent diffusion behaviors on the rumor spreading in social networks. Phys Lett A 2016;380(24):2054–63.
[13] Volterra V. Variations and fluctuations of the number of individuals in animal species living together. J Conseil 1928;3(1):3–51.
[14] Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer. Adv Eng Softw 2014;69(3):46–61.
[15] Zhao F, Zhang H, Wang L. A pareto-based discrete jaya algorithm for multiobjective carbon-efficient distributed blocking flow shop scheduling problem. IEEE Trans Ind Inf 2022.
[16] Mirjalili S, Saremi S, Mirjalili SM, et al. Multi-objective grey wolf optimizer:a novel algorithm for multi-criterion optimization. Expert Syst Appl 2015;47(4):106–19.
[17] Wang GG, Gao D, Pedrycz W. Solving multi-objective fuzzy job-shop scheduling problem by a hybrid adaptive differential evolution algorithm. IEEE Trans Ind Inf 2022.
[18] Kennedy J, Eberhart R. Particle swarm optimization, in Neural Networks. In: Proceedings, IEEE international conference on; 1995. p. 1942–8.
[19] Dandan You, Fuji Chen. Research on the prediction of network public opinion based on improved PSO and BP neural network. J Intell 2016;35(8):6.
[20] Fuji Chen, Yaju Huang. Research on the prediction of network public opinion based on SAPSORBF neural network. J Wuhan Univ Technol (Inform Manage Eng) 2017;39(04). 422-426+438.