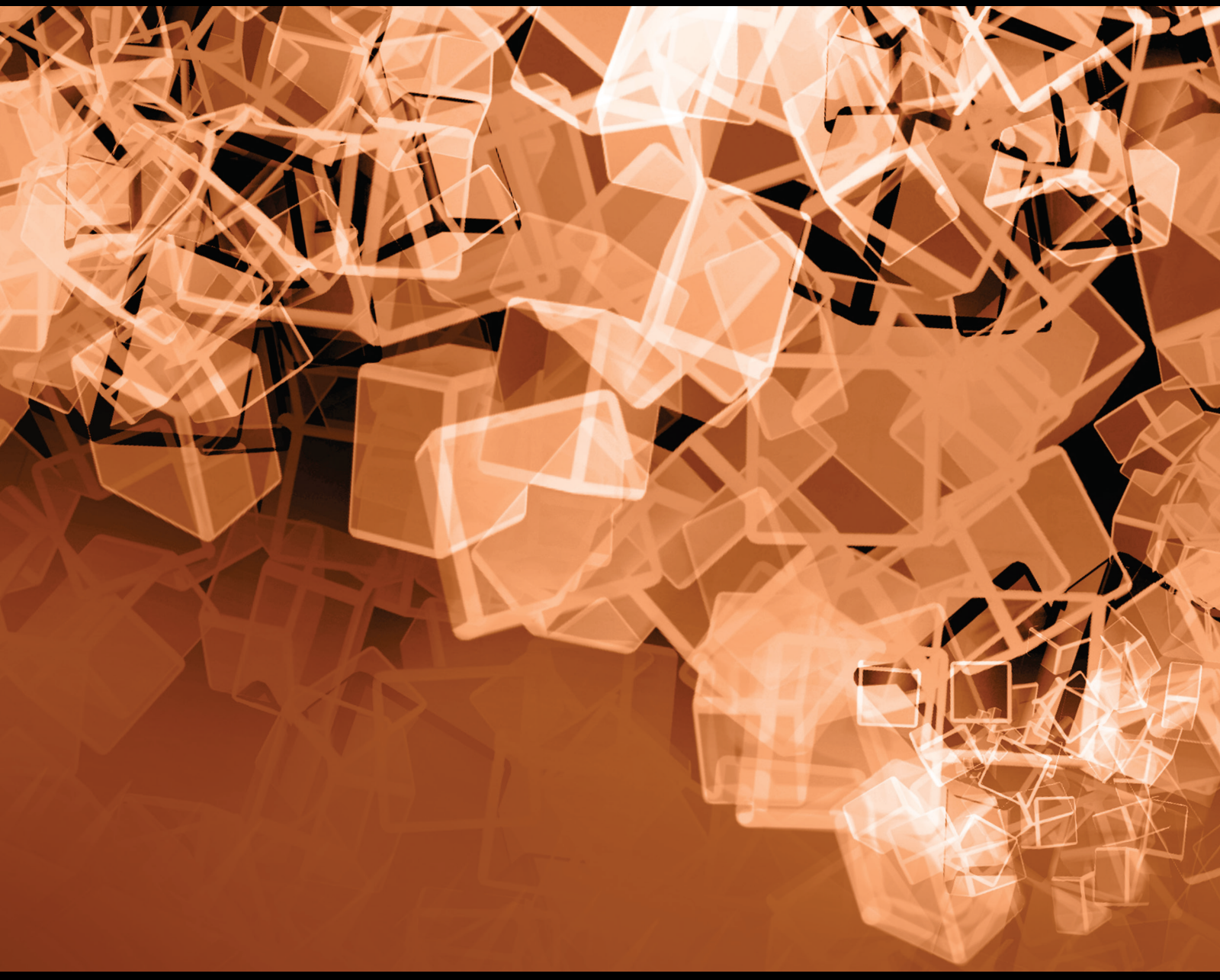
Discrete Dynamics in Nature and Society

Discrete Dynamics of Complex Interactions between Natural and Artificial Systems

Lead Guest Editor: Jorge E. Macias-Diaz
Guest Editors: Qin Sheng and Stefania Tomasiello



Discrete Dynamics of Complex Interactions between Natural and Artificial Systems

Discrete Dynamics in Nature and Society

Discrete Dynamics of Complex Interactions between Natural and Artificial Systems

Lead Guest Editor: Jorge E. Macias-Diaz
Guest Editors: Qin Sheng and Stefania Tomasiello

Contents

A Novel Regret Theory-Based Decision-Making Method Combined with the Intuitionistic Fuzzy Canberra Distance
Haiping Ren, Yunxiao Gao, and Tonghua Yang
Research Article (9 pages), Article ID 8848031, Volume 2020 (2020)

A New Theoretical Interpretation of Measurement Error and Its Uncertainty
Huisheng Shi, Xiaoming Ye, Cheng Xing, and Shijun Ding
Research Article (14 pages), Article ID 3864578, Volume 2020 (2020)

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 5530744, 19 pages
https://doi.org/10.1155/2021/5530744

Research Article

Second-Order Semi-Discretized Schemes for Solving Stochastic Quenching Models on Arbitrary Spatial Grids

Nina Garcia-Montoya,[1] Julienne Kabre,[1] Jorge E. Macías-Díaz,[2,3] and Qin Sheng[1]

[1]*Department of Mathematics and Center for Astrophysics, Space Physics and Engineering Research, Baylor University, Waco, TX 76798-7328, USA*
[2]*Department of Mathematics, School of Digital Technologies, Tallinn University, Narva Rd. 25, 10120 Tallinn, Estonia*
[3]*Departamento de Matemáticas y Física, Universidad Autónoma de Aguascalientes, Avenida Universidad 940, Ciudad Universitaria, Aguascalientes, Ags. 20131, Mexico*

Correspondence should be addressed to Qin Sheng; qin_sheng@baylor.edu

Received 23 January 2021; Revised 5 April 2021; Accepted 17 April 2021; Published 6 May 2021

Academic Editor: Chris Goodrich

Reaction-diffusion-advection equations provide precise interpretations for many important phenomena in complex interactions between natural and artificial systems. This paper studies second-order semi-discretizations for the numerical solution of reaction-diffusion-advection equations modeling quenching types of singularities occurring in numerous applications. Our investigations particularly focus at cases where nonuniform spatial grids are utilized. Detailed derivations and analysis are accomplished. Easy-to-use and highly effective second-order schemes are acquired. Computational experiments are presented to illustrate our results as well as to demonstrate the viability and capability of the new methods for solving singular quenching problems on arbitrary grid platforms.

1. Introduction

Nonlinear reaction-diffusion-advection equations have been playing an important role in interactions between natural and artificial systems. The partial differential equations provide precisely mathematical interpretations for numerous natural phenomena, such as diffusion of heat and energy, burning or quenching of fuels, energy concentrations, and transformations in different environments. Nonlinear partial differential equations also deliver popular computational tools to cell biology and cancer treatment plans [1–4].

Consider an m-dimensional heat engine, $m \geq 1$. Let u be the temperature distribution inside of its combustion chamber and u_0 be the initial temperature distribution due to the sparks. Assume that combustion occurs at the unit temperature. Further, let $0 < \sigma \leq 1$, $t > 0$, be the reciprocal chamber size index. The use of variable σ reflects the change of chamber size, which is typical when different chemical fuels are selected. Then, an idealized combustion model can be comprised as a nonlinear reaction-diffusion-advection initial-boundary value problem:

$$u_t = \sigma^2 \nabla (a \nabla u) + f(u) \mathcal{X}(u), \quad x \in \mathcal{D},\ t > 0, \tag{1}$$

$$u(x,t) = 0, \quad x \in \partial\mathcal{D},\ t > 0, \tag{2}$$

$$u(x,=0) = u_0(x), \quad x \in \mathcal{D} \cup \partial\mathcal{D}, \tag{3}$$

where $\mathcal{D} \subset \mathbb{R}^m$, $\partial\mathcal{D}$ is its boundary, ∇ is the m-dimensional gradient vector, the coefficient $a = a(x) \in C^1(\mathcal{D})$ is positive, $0 \leq u_0 \ll 1$, and $\mathcal{X}(u)$ is the stochastic indicator for $u < 1$ [5, 6].

The nonlinear source function $f(u) > 0$ is strictly increasing with respect to $0 \leq u < 1$ and

$$\begin{gathered} f(0) > 0, \\ \lim_{u \longrightarrow 1^-} f(u) = \infty. \end{gathered} \tag{4}$$

A typical example of such reaction term is

$$f(u) = \frac{1}{(1-u)^{\kappa}}, \quad \text{for } 0 \le u < 1, \tag{5}$$

where $\kappa > 0$ is the internal combustion index utilized [3, 5–7].

The solution u of (1)–(4) is said to *quench* if there exists a finite T_σ such that

$$\sup_{0<x<1} u_t(x,t) \longrightarrow \infty, \quad \text{as } t \longrightarrow T_\sigma^-. \tag{6}$$

Such a value T_σ is called the *quenching time* which is a possible combustion-reaction time [7, 8]. It has been shown that a necessary condition for quenching to occur is

$$\max_{0\le x\le 1} u(x,t) \longrightarrow 1^-, \quad \text{as } t \longrightarrow T_\sigma^-. \tag{7}$$

The study of singular reaction-diffusion-advection equations of the form (1) and (5) can be traced back to Hale and Kawarada's pioneering work [9, 10]. It was observed that when $m = \kappa = 1$ and $a \equiv 1$, there exists a *critical value* $\sigma^* \approx 0.65340$ such that the solution of (1)–(3) and (5) quenches if $\sigma < \sigma^*$. Verifications of the existence and uniqueness to more general quenching models can be found in numerous recent publications including [1, 5, 9, 11].

To solve the nonlinear differential equation problem (1)–(3) numerically is an interesting, yet challenging, multitask since we often do not know if the solution of (1)–(3) will quench or not until a proper solution procedure is taking place. In other words, in addition to approximating the solution u, a numerical method must be capable of evaluating simultaneously correct critical value σ^*, quenching time T_σ^*, quenching location x^*, and extremely sensitive and possibly unbounded derivative function u_t [5, 6, 11]. This can be difficult since values of $\max_{x\in\overline{\mathcal{D}}} u_t$ remain bounded and well-behaved until time t reaches a certain neighborhood of T^*, if it exists. Furthermore, we may also observe from properties (6) and (7) that, while $\max_{x\in\overline{\mathcal{D}}} u_t$ grows exponentially, or faster than exponentially, in the aforementioned neighborhoods, the solution u itself continuously grows but stays bounded throughout the computation until a quenching suddenly erupts [7, 9, 10]. These coexisted and very distinct characters make it extremely hard to design an effective scheme.

Due to their extremely important applications in nature and societies, numerous investigations, in both theory and computations, have been carried out for nonlinear quenching problems including (1)–(3) in recent years. Most of the existing algorithms are constructed either based on the reduced problems, that is, the stationary problems by removing the variable t, or by using fixed mesh steps (cf. [4, 7, 8, 12–14] and references therein). In these procedures, critical values, quenching times, solution u, and the rate-of-change function u_t are often approximated incorrectly. Computational procedures are less efficient and less reliable, in particular when multidimensional modeling equations are considered. Although adaptive strategies have been particularly in favor due to their great flexibility and geometric accuracy in capturing quenching singularities involved, orders of accuracies of the spatial discretization of existing adaptive methods have been limited to one due to the unpredictability of the nonuniform grids generated by adaptations [6, 8, 12].

On the other hand, effective new approaches in high-order approximations via finite difference, spectral, or finite element methods, have been introduced for solving other nonlinear partial differential equations [2, 15, 16]. Investigations have also been extended to fractional order partial differential equation problems [16, 17]. Since it is the high dimension that causes a tremendous increase in the computational cost, modern splitting techniques capable of converting higher dimensional problems into sets of single dimensional subproblems have reached a new height [13, 18]. Rigorous numerical analysis has also been given on split adaptations [18]. These have motivated our study of highly applicable and effective higher-order finite difference methods.

Note that problems (1)–(3) can be numerically treated through proper second-order stable dimensional splitting [18, 19]. The decomposition may effectively reduce the total number of operations from $\mathcal{O}(n^{3m})$ to $\mathcal{O}(n^3)$, where m is the number of dimensions and n is the number of internal mesh points used [7, 18, 20]. Thus, in this paper, we shall focus on second-order semi-discretization for a one-dimensional modeling equation with $m = 1$ and $a \equiv 1$. In the circumstance, problems (1)–(3) can be simplified to

$$u_t = \sigma^2 u_{xx} + f(u)\mathcal{X}(u), \quad 0 < x < 1, t > 0, \tag{8}$$

$$u(0,t) = u(1,t) = 0, \quad t > 0, \tag{9}$$

$$u(x,0) = u_0(x), \quad 0 \le x \le 1. \tag{10}$$

We consider an arbitrary partition of the spatial domain $[0,1]$, that is, grids

$$\Omega_n = \{0 = x_0, x_1, x_2, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n, x_{n+1} = 1\}, \tag{11}$$

for which $0 < h_k = x_{k+1} - x_k \ll 1$, $k = 0, 1, \ldots, n$. Denote $H_n = \{h_1, h_2, \ldots, h_{n+1}\}$. Apparently, such sets Ω_n, H_n are t-dependent when they are used for solving (8)–(10). However, due to the feature of semi-discretization, we prefer dropping time location indicators for the simplicity of notations. We further assume that in general $h_{k+1} \ne h_k$, $k = 0, 1, \ldots, n$. At any spatial location x_k, we adopt usual notations $\phi(x_k, t) = \phi_k(t)$ or $\phi(x_k, t) = \phi_k$ without confusion.

2. Semi-Discretization Approximations

At the time level $t > 0$, we set

$$h^{(k)} = \max\{h_{k-3}, h_{k-2}, h_{k-1}, h_k, h_{k+1}, h_{k+2}\}, \tag{12}$$

for any available index k. Without loss of generality, we drop the superindex for simplicity. We have expansions of the spatial derivatives

$$u_k'' = b_{k,-1}u_{k-1} + b_{k,0}u_k + b_{k,1}u_{k+1} + b_{k,2}u_{k+2} + b_{k,3}u_{k+3} + \mathcal{O}\left(h^2\right), \tag{13}$$

$$u_k'' = a_{k,-2}u_{k-2} + a_{k,-1}u_{k-1} + a_{k,0}u_k + a_{k,1}u_{k+1} + a_{k,2}u_{k+2} + \mathcal{O}\left(h^2\right), \tag{14}$$

$$u_k'' = c_{k,-3}u_{k-3} + c_{k,-2}u_{k-2} + c_{k,-1}u_{k-1} + c_{k,0}u_k + c_{k,1}u_{k+1} + \mathcal{O}\left(h^2\right). \tag{15}$$

Strategy 1. From the above, we must have

$$\begin{aligned} u_1'' &= b_{1,-1}u_0 + b_{1,0}u_1 + b_{1,1}u_2 + b_{1,2}u_3 + b_{1,3}u_4 + \mathcal{O}\left(h^2\right), \\ u_k'' &= a_{k,-2}u_{k-2} + a_{k,-1}u_{k-1} + a_{k,0}u_k + a_{k,1}u_{k+1} + a_{k,2}u_{k+2} + \mathcal{O}\left(h^2\right), \quad k = 2, \ldots, n-1, \\ u_n'' &= c_{n,-3}u_{n-3} + c_{n,-2}u_{n-2} + c_{n,-1}u_{n-1} + c_{n,0}u_n + c_{n,1}u_{n+1} + \mathcal{O}\left(h^2\right). \end{aligned} \tag{16}$$

Note that $u_0 = u_{n+1} = 0$ due to the homogeneous condition (9). Substituting the above into (8)–(10), we acquire that

$$\begin{aligned} u_t &= \sigma^2 Au + b + \mathcal{O}\left(h^2\right), \\ u(0) &= u_0, \end{aligned} \tag{17}$$

where

$$\begin{aligned} A &= \begin{bmatrix} b_{1,0} & b_{1,1} & b_{1,2} & b_{1,3} & & & & & \\ a_{2,-1} & a_{2,0} & a_{2,1} & a_{2,2} & & & & & \\ a_{3,-2} & a_{3,-1} & a_{3,0} & a_{3,1} & a_{3,2} & & & & \\ & a_{4,-2} & a_{4,-1} & a_{4,0} & a_{4,1} & a_{4,2} & & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & & & a_{n-2,-2} & a_{n-2,-1} & a_{n-2,0} & a_{n-2,1} & a_{n-2,2} \\ & & & & & a_{n-1,-2} & a_{n-1,-1} & a_{n-1,0} & a_{n-1,1} \\ & & & & & c_{n,-3} & c_{n,-2} & c_{n,-1} & c_{n,0} \end{bmatrix} \in \mathbb{R}^{n\times n}, \\ u &= \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n, \\ b &= \begin{bmatrix} f(u_1)\mathcal{X}(u_1) \\ f(u_2)\mathcal{X}(u_2) \\ f(u_3)\mathcal{X}(u_3) \\ \vdots \\ \vdots \\ f(u_n)\mathcal{X}(u_n) \end{bmatrix} \in \mathbb{R}^n. \end{aligned} \tag{18}$$

The band matrix A is large in size since n is usually large. But it is relatively sparse and can be handled conveniently by many existing software packages such as "sparse" subroutines in MATLAB. Dropping the truncation error term in (19), we obtain a second-order semi-discretization scheme for solving (8)–(10):

$$u_t = \sigma^2 Au + b, \tag{19}$$

$$u(0) = u_0. \tag{20}$$

Therefore, the solution of (8)–(10) can be readily approximated numerically through the solution of the nonlinear system of ordinary differential equations (19) and (20).

Strategy 2. The coefficient matrix A in (19) can be extremely stiff particularly due to the use of one-sided finite difference formulas (13) and (15). To avoid using the formulas, we may assume first that, in addition to the boundary values $u_0 = u_{n+1} = 0$, we also have values u_{-1}, u_{n+2} calculated via some other methods that are at least of second order in accuracy. In this way, we may only use (14) for our second derivative approximations. The fully centralized formulas may hopefully ease the stiffness [19]. This yields our semi-discretized new approximation

$$\begin{aligned} u_t &= \sigma^2 Bu + c + \mathcal{O}\left(h^2\right), \\ u(0) &= u_0, \end{aligned} \tag{21}$$

where

$$\begin{aligned} B &= \begin{bmatrix} a_{1,0} & a_{1,1} & a_{1,2} & & & & & & \\ a_{2,-1} & a_{2,0} & a_{2,1} & a_{2,2} & & & & & \\ a_{3,-2} & a_{3,-1} & a_{3,0} & a_{3,1} & a_{3,2} & & & & \\ & a_{4,-2} & a_{4,-1} & a_{4,0} & a_{4,1} & a_{4,2} & & & \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ & & & & a_{n-2,-2} & a_{n-2,-1} & a_{n-2,0} & a_{n-2,1} & a_{n-2,2} \\ & & & & & a_{n-1,-2} & a_{n-1,-1} & a_{n-1,0} & a_{n-1,1} \\ & & & & & & a_{n,-2} & a_{n,-1} & a_{n,0} \end{bmatrix} \in \mathbb{R}^{n\times n}, \\ u &= \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ \vdots \\ u_n \end{bmatrix} \in \mathbb{R}^n, \\ c &= \begin{bmatrix} f(u_1)\mathcal{X}(u_1) - a_{1,-2}u_{-1} \\ f(u_2)\mathcal{X}(u_2) \\ f(u_3)\mathcal{X}(u_3) \\ \vdots \\ \vdots \\ f(u_n)\mathcal{X}(u_n) - a_{n,2}u_{n+2} \end{bmatrix} \in \mathbb{R}^n. \end{aligned} \tag{22}$$

Dropping the truncation error term from the differential system, we obtain another second-order semi-discretization scheme for solving (8)–(10):

$$u_t = \sigma^2 Bu + c, \tag{23}$$

$$u(0) = u_0. \tag{24}$$

Therefore, the solution of (8)–(10) can be approximated numerically through solving the nonlinear system of ordinary differential equations (23) and (24), given that values of u_{-1}, u_{n+2} can indeed be calculated successfully.

3. Parameter Determinations

Now, let us investigate the issue and see if values u_{-1}, u_{n+2} can be computed through some direct expansions. To this end, we have

$$u_{k+1} = u_k + h_k u_k' + \frac{h_k^2}{2!}u_k'' + \frac{h_k^3}{3!}u_k''' + \frac{h_k^4}{4!}u_k^{(4)} + \frac{h_k^5}{5!}u_k^{(5)} + \frac{h_k^6}{6!}u_k^{(6)} + \mathcal{O}\left(h_k^7\right). \tag{25}$$

We observe readily that

$$\begin{aligned} u_{-1} &= u_0 - h_{-1}u_0' + \frac{h_{-1}^2}{2!}u_0'' - \frac{h_{-1}^3}{3!}u_0''' + \frac{h_{-1}^4}{4!}u_0^{(4)} - \frac{h_{-1}^5}{5!}u_0^{(5)} + \frac{h_{-1}^6}{6!}u_0^{(6)} + \mathcal{O}\left(h_{-1}^7\right), \\ u_{n+2} &= u_{n+1} + h_{n+1}u_{n+1}' + \frac{h_{n+1}^2}{2!}u_{n+1}'' + \frac{h_{n+1}^3}{3!}u_{n+1}''' + \frac{h_{n+1}^4}{4!}u_{n+1}^{(4)} + \frac{h_{n+1}^5}{5!}u_{n+1}^{(5)} + \frac{h_{n+1}^6}{6!}u_{n+1}^{(6)} + \mathcal{O}\left(h_{n+1}^7\right). \end{aligned} \tag{26}$$

For the simplicity in calculations, we may set $h_{-1} = h_0, h_{n+1} = h_n$. Thus,

$$\begin{aligned} u_{-1} &= u_0 - h_0 u_0' + \frac{h_0^2}{2!}u_0'' - \frac{h_0^3}{3!}u_0''' + \frac{h_0^4}{4!}u_0^{(4)} - \frac{h_0^5}{5!}u_0^{(5)} + \frac{h_0^6}{6!}u_0^{(6)} + \mathcal{O}\left(h_0^7\right), \\ u_{n+2} &= u_n + h_n u_n' + \frac{h_n^2}{2!}u_n'' + \frac{h_n^3}{3!}u_n''' + \frac{h_n^4}{4!}u_n^{(4)} + \frac{h_n^5}{5!}u_n^{(5)} + \frac{h_n^6}{6!}u_n^{(6)} + \mathcal{O}\left(h_n^7\right). \end{aligned} \tag{27}$$

To use the above two formulations for a quadratic order approximation, we need to balance coefficients related to u_0', u_0'' and u_n', u_n'' at internal spatial mesh points. Recall (13) and (15) and set

$$u_0' = \beta_{1,0}u_0 + \beta_{1,1}u_1 + \beta_{1,2}u_2 + \mathcal{O}\left(h^2\right), \tag{28}$$

$$u_0'' = \beta_{2,0}u_0 + \beta_{2,1}u_1 + \beta_{2,2}u_2 + \beta_{2,3}u_3 + \beta_{2,4}u_4 + \mathcal{O}\left(h^2\right), \tag{29}$$

$$u_n' = \gamma_{1,n-2}u_{n-2} + \gamma_{1,n-1}u_{n-1} + \gamma_{1,n}u_n + \mathcal{O}\left(h^2\right), \tag{30}$$

$$u_n'' = \gamma_{2,n-4}u_{n-4} + \gamma_{2,n-3}u_{n-3} + \gamma_{2,n-2}u_{n-2} + \gamma_{2,n-1}u_{n-1} + \gamma_{2,n}u_n + \mathcal{O}\left(h^2\right). \tag{31}$$

Recall (25). To find the coefficients $\beta_{1,0}, \beta_{1,1}, \beta_{1,2}$, we let

$$\begin{aligned} u_0' = {} & \beta_{1,0}u_0 + \beta_{1,1}\left[u_0 + h_0u_0' + \frac{h_0^2}{2!}u_0'' + \frac{h_0^3}{3!}u_0''' + \frac{h_0^4}{4!}u_0^{(4)} + \frac{h_0^5}{5!}u_0^{(5)} + \frac{h_0^6}{6!}u_0^{(6)} + \cdots\right] \\ & + \beta_{1,2}\left[u_0 + (h_0+h_1)u_0' + \frac{(h_0+h_1)^2}{2!}u_0'' + \frac{(h_0+h_1)^3}{3!}u_0' + \frac{(h_0+h_1)^4}{4!}u_0^{(4)} + \frac{(h_0+h_1)^5}{5!}u_0^{(5)} + \frac{(h_0+h_1)^6}{6!}u_0^{(6)} + \cdots\right] \\ & + \mathcal{O}\left(h^2\right). \end{aligned} \tag{32}$$

It follows therefore

$$\begin{aligned} \beta_{1,0} + \beta_{1,1} + \beta_{1,2} &= 0, \\ \beta_{1,1} h_0 + \beta_{1,2}(h_0 + h_1) &= 1, \\ \beta_{1,1} h_0^2 + \beta_{1,2}(h_0 + h_1)^2 &= 0. \end{aligned} \tag{33}$$

It turns out that

$$\begin{aligned} \beta_{1,0} &= -\beta_{1,1} - \beta_{1,2}, \\ 1 &= \beta_{1,1} h_0 + \beta_{1,2}(h_0 + h_1), \\ \beta_{1,1} &= -\frac{(h_0 + h_1)^2}{h_0^2}\beta_{1,2}. \end{aligned} \tag{34}$$

From the last two equations,

$$\begin{aligned} &-\frac{(h_0 + h_1)^2}{h_0^2}\beta_{1,2} h_0 + \beta_{1,2}(h_0 + h_1) \\ &= \left(h_0 + h_1 - \frac{(h_0 + h_1)^2}{h_0}\right)\beta_{1,2} = 1. \end{aligned} \tag{35}$$

Hence,

$$\begin{aligned} \beta_{1,1} &= \frac{h_0 + h_1}{h_0 h_1}, \\ \beta_{1,2} &= -\frac{h_0}{(h_0 + h_1)h_1}. \end{aligned} \tag{36}$$

Consequently,

$$\begin{aligned} \beta_{1,0} &= -\frac{h_0 + h_1}{h_0 h_1} + \frac{h_0}{(h_0 + h_1)h_1} = \frac{1}{h_1}\left(\frac{h_0^2 - (h_0 + h_1)^2}{h_0(h_0 + h_1)}\right) \\ &= -\frac{2h_0 + h_1}{h_0(h_0 + h_1)}. \end{aligned} \tag{37}$$

Similarly, to find coefficients $\gamma_{1,n-2}, \gamma_{1,n-1}, \gamma_{1,n}$, we consider

$$\begin{aligned} u_n' = \gamma_{1,n-2}\Bigg[& u_n - (h_{n-2} + h_{n-1})u_n' + \frac{(h_{n-2} + h_{n-1})^2}{2!}u_n'' - \frac{(h_{n-2} + h_{n-1})^3}{3!}u_n'' + \frac{(h_{n-2} + h_{n-1})^4}{4!}u_n^{(4)} \\ & - \frac{(h_{n-2} + h_{n-1})^5}{5!}u_n^{(5)} + \frac{(h_{n-2} + h_{n-1})^6}{6!}u_n^{(6)} + \cdots \Bigg] \\ & + \gamma_{1,n-1}\left[u_n - h_{n-1}u_0' + \frac{h_{n-1}^2}{2!}u_n'' - \frac{h_{n-1}^3}{3!}u_n'' + \frac{h_{n-1}^4}{4!}u_n^{(4)} - \frac{h_{n-1}^5}{5!}u_n^{(5)} + \frac{h_{n-1}^6}{6!}u_n^{(6)} + \cdots\right] + \gamma_{1,n}u_n + \mathcal{O}(h^2). \end{aligned} \tag{38}$$

Balancing the coefficients, we have

$$\begin{aligned} \gamma_{1,n-2} + \gamma_{1,n-1} + \gamma_{1,n} &= 0, \\ -\gamma_{1,n-2}(h_{n-2} + h_{n-1}) - \gamma_{1,n-1}h_{n-1} &= 1, \\ \gamma_{1,n-2}(h_{n-2} + h_{n-1})^2 + \gamma_{1,n-1}h_{n-1}^2 &= 0. \end{aligned} \tag{39}$$

It follows subsequently that

$$\begin{aligned} \gamma_{1,n} &= -\gamma_{1,n-2} - \gamma_{1,n-1}, \\ -1 &= \gamma_{1,n-2}(h_{n-2} + h_{n-1}) + \gamma_{1,n-1}h_{n-1}, \\ \gamma_{1,n-1} &= -\frac{(h_{n-2} + h_{n-1})^2}{h_{n-1}^2}\gamma_{1,n-2}. \end{aligned} \tag{40}$$

From the last two equations,

$$\gamma_{1,n-2}(h_{n-2} + h_{n-1}) - \frac{(h_{n-2} + h_{n-1})^2}{h_{n-1}^2}\gamma_{1,n-2}h_{n-1} = -\frac{(h_{n-2} + h_{n-1})h_{n-2}}{h_{n-1}}\gamma_{1,n-2} = -1. \tag{41}$$

Hence,

$$\begin{aligned} \gamma_{1,n-2} &= \frac{h_{n-1}}{(h_{n-2} + h_{n-1})h_{n-2}}, \\ \gamma_{1,n-1} &= -\frac{h_{n-2} + h_{n-1}}{h_{n-2}h_{n-1}}, \end{aligned} \tag{42}$$

and furthermore,

$$\begin{aligned} \gamma_{1,n} &= -\frac{h_{n-1}}{(h_{n-2} + h_{n-1})h_{n-2}} + \frac{h_{n-2} + h_{n-1}}{h_{n-2}h_{n-1}} \\ &= \frac{h_{n-2} + 2h_{n-1}}{(h_{n-2} + h_{n-1})h_{n-1}}. \end{aligned} \tag{43}$$

Now, recall the expansion (25). To determine coefficients $\beta_{2,0}, \beta_{2,1}, \beta_{2,2}, \beta_{2,3}, \beta_{2,4}$ in (29), we have

$$\begin{aligned}
u_0'' = {} & \beta_{2,0}u_0 + \beta_{2,1}\left[u_0 + h_0u_0' + \frac{h_0^2}{2!}u_0'' + \frac{h_0^3}{3!}u_0'' + \frac{h_0^4}{4!}u_0^{(4)} + \frac{h_0^5}{5!}u_0^{(5)} + \frac{h_0^6}{6!}u_0^{(6)} + \cdots\right] \\
& + \beta_{2,2}\left[u_0 + (h_0 + h_1)u_0' + \frac{(h_0 + h_1)^2}{2!}u_0'' + \frac{(h_0 + h_1)^3}{3!}u_0'' + \frac{(h_0 + h_1)^4}{4!}u_0^{(4)}\right. \\
& \left. + \frac{(h_0 + h_1)^5}{5!}u_0^{(5)} + \frac{(h_0 + h_1)^6}{6!}u_0^{(6)} + \cdots\right] \\
& + \beta_{2,3}\left[u_0 + (h_0 + h_1 + h_2)u_0' + \frac{(h_0 + h_1 + h_2)^2}{2!}u_0'' + \frac{(h_0 + h_1 + h_2)^3}{3!}u_0''\right. \\
& \left. + \frac{(h_0 + h_1 + h_2)^4}{4!}u_0^{(4)} + \frac{(h_0 + h_1 + h_2)^5}{5!}u_0^{(5)} + \frac{(h_0 + h_1 + h_2)^6}{6!}u_0^{(6)} + \cdots\right] \\
& + \beta_{2,4}\left[u_0 + (h_0 + h_1 + h_2 + h_3)u_0' + \frac{(h_0 + h_1 + h_2 + h_3)^2}{2!}u_0'' + \frac{(h_0 + h_1 + h_2 + h_3)^3}{3!}u_0''\right. \\
& \left. + \frac{(h_0 + h_1 + h_2 + h_3)^4}{4!}u_0^{(4)} + \frac{(h_0 + h_1 + h_2 + h_3)^5}{5!}u_0^{(5)} + \frac{(h_0 + h_1 + h_2 + h_3)^6}{6!}u_0^{(6)} + \cdots\right] + \mathcal{O}\left(h^2\right).
\end{aligned} \tag{44}$$

From the above, we obtain the following linear system:

$$\begin{aligned}
\beta_{2,0} + \beta_{2,1} + \beta_{2,2} + \beta_{2,3} + \beta_{2,4} &= 0, \\
\beta_{2,1}h_0 + \beta_{2,2}(h_0 + h_1) + \beta_{2,3}(h_0 + h_1 + h_2) + \beta_{2,4}(h_0 + h_1 + h_2 + h_3) &= 0, \\
\beta_{2,1}h_0^2 + \beta_{2,2}(h_0 + h_1)^2 + \beta_{2,3}(h_0 + h_1 + h_2)^2 + \beta_{2,4}(h_0 + h_1 + h_2 + h_3)^2 &= 2, \\
\beta_{2,1}h_0^3 + \beta_{2,2}(h_0 + h_1)^3 + \beta_{2,3}(h_0 + h_1 + h_2)^3 + \beta_{2,4}(h_0 + h_1 + h_2 + h_3)^3 &= 0, \\
\beta_{2,1}h_0^4 + \beta_{2,2}(h_0 + h_1)^4 + \beta_{2,3}(h_0 + h_1 + h_2)^4 + \beta_{2,4}(h_0 + h_1 + h_2 + h_3)^4 &= 0.
\end{aligned} \tag{45}$$

The system can be conveniently solved digitally via any existing solution package. By the same token, for evaluating coefficients $\gamma_{2,n-4}, \gamma_{2,n-3}, \gamma_{2,n-2}, \gamma_{2,n-1}, \gamma_{2,n}$, in (31), we find that

$$
\begin{aligned}
u_n'' = {} & +\gamma_{2,n-4}\Bigg[u_n - (h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})u_n' + \frac{(h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})^2}{2!}u_n'' \\
& -\frac{(h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})^3}{3!}u_n''' + \frac{(h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})^4}{4!}u_n^{(4)} \\
& -\frac{(h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})^5}{5!}u_n^{(5)} + \frac{(h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})^6}{6!}u_n^{(6)} + \cdots\Bigg] \\
& +\gamma_{2,n-3}\Bigg[u_n - (h_{n-3}+h_{n-2}+h_{n-1})u_n' + \frac{(h_{n-3}+h_{n-2}+h_{n-1})^2}{2!}u_n'' - \frac{(h_{n-3}+h_{n-2}+h_{n-1})^3}{3!}u_n''' \\
& +\frac{(h_{n-3}+h_{n-2}+h_{n-1})^4}{4!}u_n^{(4)} - \frac{(h_{n-3}+h_{n-2}+h_{n-1})^5}{5!}u_n^{(5)} + \frac{(h_{n-3}+h_{n-2}+h_{n-1})^6}{6!}u_n^{(6)} + \cdots\Bigg] \\
& +\gamma_{2,n-2}\Bigg[u_n - (h_{n-2}+h_{n-1})u_n' + \frac{(h_{n-2}+h_{n-1})^2}{2!}u_n'' - \frac{(h_{n-2}+h_{n-1})^3}{3!}u_n''' \\
& +\frac{(h_{n-2}+h_{n-1})^4}{4!}u_n^{(4)} - \frac{(h_{n-2}+h_{n-1})^5}{5!}u_n^{(5)} + \frac{(h_{n-2}+h_{n-1})^6}{6!}u_n^{(6)} + \cdots\Bigg] \\
& +\gamma_{2,n-1}\Bigg[u_n - h_{n-1}u_0' + \frac{h_{n-1}^2}{2!}u_n'' - \frac{h_{n-1}^3}{3!}u_n''' + \frac{h_{n-1}^4}{4!}u_n^{(4)} - \frac{h_{n-1}^5}{5!}u_n^{(5)} + \frac{h_{n-1}^6}{6!}u_n^{(6)} + \cdots\Bigg] + \gamma_{2,n}u_n + \mathcal{O}(h^2).
\end{aligned}
\tag{46}
$$

The above expansion leads to the following non-homogeneous linear system:

$$
\begin{aligned}
\gamma_{2,n-4}+\gamma_{2,n-3}+\gamma_{2,n-2}+\gamma_{2,n-1}+\gamma_{2,n} &= 0, \\
\gamma_{2,n-4}(h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})+\gamma_{2,n-3}(h_{n-3}+h_{n-2}+h_{n-1})+\gamma_{2,n-2}(h_{n-2}+h_{n-1})+\gamma_{2,n-1}h_{n-1} &= 0, \\
\gamma_{2,n-4}(h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})^2+\gamma_{2,n-3}(h_{n-3}+h_{n-2}+h_{n-1})^2+\gamma_{2,n-2}(h_{n-2}+h_{n-1})^2+\gamma_{2,n-1}h_{n-1}^2 &= 2, \\
\gamma_{2,n-4}(h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})^3+\gamma_{2,n-3}(h_{n-3}+h_{n-2}+h_{n-1})^3+\gamma_{2,n-2}(h_{n-2}+h_{n-1})^3+\gamma_{2,n-1}h_{n-1}^3 &= 0, \\
\gamma_{2,n-4}(h_{n-4}+h_{n-3}+h_{n-2}+h_{n-1})^4+\gamma_{2,n-3}(h_{n-3}+h_{n-2}+h_{n-1})^4+\gamma_{2,n-2}(h_{n-2}+h_{n-1})^4+\gamma_{2,n-1}h_{n-1}^4 &= 0.
\end{aligned}
\tag{47}
$$

Now, recall (14). To determine coefficients $a_{k-2}, a_{k-1}, a_k, a_{k+1}, a_{k+2}$, we have

$$
\begin{aligned}
u_k'' = {} & [a_{k-2} + a_{k-1} + a_k + a_{k+1} + a_{k+2}]u_k \\
& + [-(h_{k-2} + h_{k-1})a_{k-2} - h_{k-1}a_{k-1} + h_k a_{k+1} + (h_k + h_{k+1})a_{k+2}]u_k' \\
& + \left[\frac{(h_{k-2} + h_{k-1})^2}{2!}a_{k-2} + \frac{h_{k-1}^2}{2!}a_{k-1} + \frac{h_k^2}{2!}a_{k+1} + \frac{(h_k + h_{k+1})^2}{2!}a_{k+2}\right]u_k'' \\
& + \left[-\frac{(h_{k-2} + h_{k-1})^3}{3!}a_{k-2} - \frac{h_{k-1}^3}{3!}a_{k-1} + \frac{h_k^3}{3!}a_{k+1} + \frac{(h_k + h_{k+1})^3}{3!}a_{k+2}\right]u_k'' \\
& + \left[\frac{(h_{k-2} + h_{k-1})^4}{4!}a_{k-2} + \frac{h_{k-1}^4}{4!}a_{k-1} + \frac{h_k^4}{4!}a_{k+1} + \frac{(h_k + h_{k+1})^4}{4!}a_{k+2}\right]u_k^{(4)} \\
& + \left[-\frac{(h_{k-2} + h_{k-1})^5}{5!}a_{k-2} - \frac{h_{k-1}^5}{5!}a_{k-1} + \frac{h_k^5}{5!}a_{k+1} + \frac{(h_k + h_{k+1})^5}{5!}a_{k+2}\right]u_k^{(5)} \\
& + \left[\frac{(h_{k-2} + h_{k-1})^6}{6!}a_{k-2} + \frac{h_{k-1}^6}{6!}a_{k-1} + \frac{h_k^6}{6!}a_{k+1} + \frac{(h_k + h_{k+1})^6}{6!}a_{k+2}\right]u_k^{(6)} + \mathcal{O}(h^7).
\end{aligned}
\tag{48}
$$

Therefore, to guarantee overall *second-order approximations* of the required function values for $k = 1, 2, \ldots, n$, we must ask that

$$
\begin{aligned}
a_{k-2} + a_{k-1} + a_k + a_{k+1} + a_{k+2} &= 0, \\
-(h_{k-2} + h_{k-1})a_{k-2} - h_{k-1}a_{k-1} + h_k a_{k+1} + (h_k + h_{k+1})a_{k+2} &= 0, \\
\frac{(h_{k-2} + h_{k-1})^2}{2!}a_{k-2} + \frac{h_{k-1}^2}{2!}a_{k-1} + \frac{h_k^2}{2!}a_{k+1} + \frac{(h_k + h_{k+1})^2}{2!}a_{k+2} &= 1, \\
-\frac{(h_{k-2} + h_{k-1})^3}{3!}a_{k-2} - \frac{h_{k-1}^3}{3!}a_{k-1} + \frac{h_k^3}{3!}a_{k+1} + \frac{(h_k + h_{k+1})^3}{3!}a_{k+2} &= 0, \\
\frac{(h_{k-2} + h_{k-1})^4}{4!}a_{k-2} + \frac{h_{k-1}^4}{4!}a_{k-1} + \frac{h_k^4}{4!}a_{k+1} + \frac{(h_k + h_{k+1})^4}{4!}a_{k+2} &= 0,
\end{aligned}
\tag{49}
$$

which can be compressed into a matrix form

$$
M_k a = f, \tag{50}
$$

where

$$
\begin{aligned}
M_k &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -h_{k-2}-h_{k-1} & -h_{k-1} & 0 & h_k & h_k+h_{k+1} \\ (h_{k-2}+h_{k-1})^2 & h_{k-1}^2 & 0 & h_k^2 & (h_k+h_{k+1})^2 \\ -(h_{k-2}+h_{k-1})^3 & -h_{k-1}^3 & 0 & h_k^3 & (h_k+h_{k+1})^3 \\ (h_{k-2}+h_{k-1})^4 & h_{k-1}^4 & 0 & h_k^4 & (h_k+h_{k+1})^4 \end{bmatrix}, \\
a &= \begin{bmatrix} a_{k-2} \\ a_{k-1} \\ a_k \\ a_{k+1} \\ a_{k+2} \end{bmatrix}, \\
f &= \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}, \quad k = 1, 2, \ldots, n.
\end{aligned}
\tag{51}
$$

Furthermore, for (13), we observe that

$$
\begin{aligned}
u_k'' &= b_{k-1}\left[u_k - h_{k-1}u_k' + \frac{h_{k-1}^2}{2!}u_k'' - \frac{h_{k-1}^3}{3!}u_k''' + \frac{h_{k-1}^4}{4!}u_k^{(4)} - \frac{h_{k-1}^5}{5!}u_k^{(5)} + \frac{h_{k-1}^6}{6!}u_k^{(6)}\right] + b_k u_k \\
&\quad + b_{k+1}\left[u_k + h_k u_k' + \frac{h_k^2}{2!}u_k'' + \frac{h_k^3}{3!}u_k''' + \frac{h_k^4}{4!}u_k^{(4)} + \frac{h_k^5}{5!}u_k^{(5)} + \frac{h_k^6}{6!}u_k^{(6)}\right] \\
&\quad + b_{k+2}\left[\begin{array}{l} u_k + (h_k+h_{k+1})u_k' + \dfrac{(h_k+h_{k+1})^2}{2!}u_k'' + \dfrac{(h_k+h_{k+1})^3}{3!}u_k''' \\ + \dfrac{(h_k+h_{k+1})^4}{4!}u_k^{(4)} + \dfrac{(h_k+h_{k+1})^5}{5!}u_k^{(5)} + \dfrac{(h_k+h_{k+1})^6}{6!}u_k^{(6)} \end{array}\right] \\
&\quad + b_{k+3}\left[\begin{array}{l} u_k + (h_k+h_{k+1}+h_{k+2})u_k' + \dfrac{(h_k+h_{k+1}+h_{k+2})^2}{2!}u_k'' + \dfrac{(h_k+h_{k+1}+h_{k+2})^3}{3!}u_k''' \\ + \dfrac{(h_k+h_{k+1}+h_{k+2})^4}{4!}u_k^{(4)} + \dfrac{(h_k+h_{k+1}+h_{k+2})^5}{5!}u_k^{(5)} + \dfrac{(h_k+h_{k+1}+h_{k+2})^6}{6!}u_k^{(6)} \end{array}\right] + \mathcal{O}(h^7) \\
&= [b_{k-1}+b_k+b_{k+1}+b_{k+2}+b_{k+3}]u_k \\
&\quad + [-h_{k-1}b_{k-1} + h_k b_{k+1} + (h_k+h_{k+1})b_{k+2} + (h_k+h_{k+1}+h_{k+2})b_{k+3}]u_k' \\
&\quad + \left[\frac{h_{k-1}^2}{2!}b_{k-1} + \frac{h_k^2}{2!}b_{k+1} + \frac{(h_k+h_{k+1})^2}{2!}b_{k+2} + \frac{(h_k+h_{k+1}+h_{k+2})^2}{2!}b_{k+3}\right]u_k'' \\
&\quad + \left[-\frac{h_{k-1}^3}{3!}b_{k-1} + \frac{h_k^3}{3!}b_{k+1} + \frac{(h_k+h_{k+1})^3}{3!}b_{k+2} + \frac{(h_k+h_{k+1}+h_{k+2})^3}{3!}b_{k+3}\right]u_k''' \\
&\quad + \left[\frac{h_{k-1}^4}{4!}b_{k-1} + \frac{h_k^4}{4!}b_{k+1} + \frac{(h_k+h_{k+1})^4}{4!}b_{k+2} + \frac{(h_k+h_{k+1}+h_{k+2})^4}{4!}b_{k+3}\right]u_k^{(4)} \\
&\quad + \left[-\frac{h_{k-1}^5}{5!}b_{k-1} + \frac{h_k^5}{5!}b_{k+1} + \frac{(h_k+h_{k+1})^5}{5!}b_{k+2} + \frac{(h_k+h_{k+1}+h_{k+2})^5}{5!}b_{k+3}\right]u_k^{(5)} \\
&\quad + \left[\frac{h_{k-1}^6}{6!}b_{k-1} + \frac{h_k^6}{6!}b_{k+1} + \frac{(h_k+h_{k+1})^6}{6!}b_{k+2} + \frac{(h_k+h_{k+1}+h_{k+2})^6}{6!}b_{k+3}\right]u_k^{(6)} + \mathcal{O}(h^7).
\end{aligned}
\tag{52}
$$

Therefore, the result is as follows:

$$\begin{aligned}
b_{k-1} + b_k + b_{k+1} + b_{k+2} + b_{k+3} &= 0,\\
-h_{k-1}b_{k-1} + h_k b_{k+1} + (h_k + h_{k+1})b_{k+2} + (h_k + h_{k+1} + h_{k+2})b_{k+3} &= 0,\\
h_{k-1}^2 b_{k-1} + h_k^2 b_{k+1} + (h_k + h_{k+1})^2 b_{k+2} + (h_k + h_{k+1} + h_{k+2})^2 b_{k+3} &= 2,\\
-h_{k-1}^3 b_{k-1} + h_k^3 b_{k+1} + (h_k + h_{k+1})^3 b_{k+2} + (h_k + h_{k+1} + h_{k+2})^3 b_{k+3} &= 0,\\
h_{k-1}^4 b_{k-1} + h_k^4 b_{k+1} + (h_k + h_{k+1})^4 b_{k+2} + (h_k + h_{k+1} + h_{k+2})^4 b_{k+3} &= 0.
\end{aligned} \tag{53}$$

It can be rewritten as

$$M_l b = g, \tag{54}$$

where

$$M_l = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -h_{k-1} & 0 & h_k & h_k + h_{k+1} & h_k + h_{k+1} + h_{k+2} \\ h_{k-1}^2 & 0 & h_k^2 & (h_k + h_{k+1})^2 & (h_k + h_{k+1} + h_{k+2})^2 \\ -h_{k-1}^3 & 0 & h_k^3 & (h_k + h_{k+1})^3 & (h_k + h_{k+1} + h_{k+2})^3 \\ h_{k-1}^4 & 0 & h_k^4 & (h_k + h_{k+1})^4 & (h_k + h_{k+1} + h_{k+2})^4 \end{bmatrix},$$

$$b = \begin{bmatrix} b_{k-1} \\ b_k \\ b_{k+1} \\ b_{k+2} \\ b_{k+3} \end{bmatrix}, \tag{55}$$

$$g = \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}.$$

Now, for (15), we have

$$
\begin{aligned}
u_k' = {} & [c_{k-3} + c_{k-2} + c_{k-1} + c_k + c_{k+1}]u_k + \begin{bmatrix} -(h_{k-3} + h_{k-2} + h_{k-1})c_{k-3} \\ -(h_{k-2} + h_{k-1})c_{k-2} - h_{k-1}c_{k-1} + h_k c_{k+1} \end{bmatrix} u_k' \\
& + \left[\frac{(h_{k-3} + h_{k-2} + h_{k-1})^2}{2!} c_{k-3} + \frac{(h_{k-2} + h_{k-1})^2}{2!} c_{k-2} + \frac{h_{k-1}^2}{2!} c_{k-1} + \frac{h_k^2}{2!} c_{k+1} \right] u_k'' \\
& + \left[-\frac{(h_{k-3} + h_{k-2} + h_{k-1})^3}{3!} c_{k-3} - \frac{(h_{k-2} + h_{k-1})^3}{3!} c_{k-2} - \frac{h_{k-1}^3}{3!} c_{k-1} + \frac{h_k^3}{3!} c_{k+1} \right] u_k''' \\
& + \left[\frac{(h_{k-3} + h_{k-2} + h_{k-1})^4}{4!} c_{k-3} + \frac{(h_{k-2} + h_{k-1})^4}{4!} c_{k-2} + \frac{h_{k-1}^4}{4!} c_{k-1} + \frac{h_k^4}{4!} c_{k+1} \right] u_k^{(4)} \\
& + \left[\frac{(h_{k-3} + h_{k-2} + h_{k-1})^5}{5!} c_{k-3} - \frac{(h_{k-2} + h_{k-1})^5}{5!} c_{k-2} - \frac{h_{k-1}^5}{5!} c_{k-1} + \frac{h_k^5}{5!} c_{k+1} \right] u_k^{(5)} \\
& + \left[\frac{(h_{k-3} + h_{k-2} + h_{k-1})^6}{6!} c_{k-3} + \frac{(h_{k-2} + h_{k-1})^6}{6!} c_{k-2} + \frac{h_{k-1}^6}{6!} c_{k-1} + \frac{h_k^6}{6!} c_{k+1} \right] u_k^{(6)} + \mathcal{O}(h^7).
\end{aligned}
\tag{56}
$$

This gives us the linear system

$$M_r c = r, \tag{57}$$

where

$$
\begin{aligned}
M_r &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ -h_{k-3} - h_{k-2} - h_{k-1} & -h_{k-2} - h_{k-1} & -h_{k-1} & 0 & h_k \\ (h_{k-3} + h_{k-2} + h_{k-1})^2 & (h_{k-2} + h_{k-1})^2 & h_{k-1}^2 & 0 & h_k^2 \\ -(h_{k-3} + h_{k-2} + h_{k-1})^3 & -(h_{k-2} + h_{k-1})^3 & -h_{k-1}^3 & 0 & h_k^3 \\ (h_{k-3} + h_{k-2} + h_{k-1})^4 & (h_{k-2} + h_{k-1})^4 & h_{k-1}^4 & 0 & h_k^4 \end{bmatrix}, \\
c &= \begin{bmatrix} c_{k-3} \\ c_{k-2} \\ c_{k-1} \\ c_k \\ c_{k+1} \end{bmatrix}, \\
r &= \begin{bmatrix} 0 \\ 0 \\ 2 \\ 0 \\ 0 \end{bmatrix}.
\end{aligned}
\tag{58}
$$

Based on the above investigations, we acquire the following result.

Theorem 1. *If matrices M_k, M_l, and M_r are nonsingular based on the selection of H_n, then the coefficient B in (23) is uniquely determined.*

We note that properties, such as the spectrums, of matrices M_k, M_l, M_r and A, B still remain to be studied based on arbitrary sets of H_n at each temporal level of t. Bear in mind that nonuniform mesh steps in H_n are determined through particular adaptive procedure employed, such as the arc-length monitoring functions investigated in [7, 13, 17, 19].

4. Simulation Experiments

Let $m = 2$ and $\kappa = 1$. We consider two typical two-dimensional reaction-diffusion-advection modeling problems of (1)–(3) associated with (5).

Example 1. We are interested in natural quenching situations [21] with $\sigma = \sqrt{5}$ and

$$a(x, y) = \exp\left\{-10\left[\left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2\right]\right\}, \quad 0 \le x, y \le 1,$$

$$u_0(x, y) = \frac{1}{40}\left[(1 - \cos(2\pi x))(1 - \cos(4\pi y))\right], \quad 0 \le x, y \le 1. \tag{59}$$

We further set $\mathcal{X}(u) \equiv 1$ for the simplicity in illustrations. This problem can be viewed as a simplification of the internal combustion when dual ignition sparks are utilized [2, 3]. The partial differential equation problem is first split to one-dimensional subproblems via exponential splitting and then Strategies 1 and 2 are applied, respectively [8, 14, 18].

The semi-discretization mesh density controller is chosen to be $n = 200$, while the temporal discretization parameter $\tau = \lambda h$ is fixed, where $\lambda = (1/10) > 0$ is the Courant-Friedrichs-Lewy constant, and $h = \max_k h_k$ is taken in an appropriate temporal level, since temporal adaptations are not a goal of the current study. Standard exponentially evolving grid adaption is adopted [6, 12]. Cubic splines are used for passing data on each temporal level since spatial grids are moved. To see more clearly solution profiles, surface plots are only shown on internal mesh points without the fixed homogeneous boundary data (2).

It is observed that, due to the strong smoothness effect of the diffusion operator, the initial temperature distribution u_0 due to the dual ignition quickly converges to a smooth surface as time increases. The surface further forms a symmetric distribution with respect to the center of the spatial domain. Effects of the two sparks become hardly noticeable. The phenomenon is demonstrated in Figure 1 at the second, 10,000th, 20,000th, and 30,000th temporal levels, respectively. It has also been found that the numerical solution is monotonically increasing pointwise as t increases.

Another extremely interesting physical function to monitor in quenching-combustion situations is the temporal derivative, or rate of change function, u_t [7, 21]. The evolution of u_t accelerates rapidly as time t is approaching the quenching time T^*, which is approximately $69727\tau \approx 0.47624479$. To see more details, we show the rate of change function u_t in four temporal levels: $T = 69600\tau, 69700\tau, 69720\tau$, and 69726τ under the same scale in Figure 2. A rapid increase of the peak values can be observed. These consistently match the latest observations reported in [7, 21].

To see more precisely profiles of u and u_t, we plot evolutional trajectories of the maximal values and mean values of them for $T \in (0, T^*)$ in Figure 3. It can be seen that the maximum of the temperature field derivative u_t increases exponentially as $t \longrightarrow T^*$, while the solution quenches peacefully. The rapid increase of averaging temperature in the combustion chamber is particularly important to engineering applications, since it indicates the success of the quick release of the chemical energy from fuel [3]. A y-semi-logarithmic scale is used to show greater details of the phenomenon, especially for the derivative function $\max_{0\le x,y\le 1} u_t(x, y, T)$.

In Figure 4, three-dimensional surfaces of u and u_t, together with their x-(u/u_t) plane projections, immediately before the quench are given. It can be seen that while the peak temperature, which is at the center of the combustion chamber, gradually approaches the fuel ignition, distribution of u tends to occupy the entire space of the chamber. Original influence of the dual sparks disappears. On the other hand, however, rates of the temperature change, u_t, increase exponentially to the infinity at the center. The semi-discretized scheme is highly stable and captures the phenomenon satisfactorily in about seventy thousand time steps.

Example 2. As a highly realistic extension, we consider another favorable natural quenching case with $\sigma = \sqrt{5}$ and a much simplified four ignition spark initial condition [2, 3]:

$$u_0(x, y) = \frac{1}{40}\left[(1 - \cos(4\pi x))(1 - \cos(4\pi y))\right], \quad 0 \le x, y \le 1. \tag{60}$$

We continue setting $\mathcal{X}(u) \equiv 1$. Again, we consider corresponding one-dimensional subproblems via exponential splitting and then adopting Strategies 1 and 2, respectively. The same spatial mesh density controller and Courant-Friedrichs-Lewy constant are used for most of the computations until t is extremely close to the quenching time. Standard exponentially evolving grid adaption is adopted in the final stages of solution calculations [12]. Cubic splines are again used for passing data on each temporal level since spatial grids are moved. We only plot solutions on internal mesh points without showing the homogeneous boundary data (2).

Figure 5 is devoted to four key moments of the solution procedures. Distributions of the temperature u and its temporal derivative u_t inside the rectangular combustion chamber are simulated. The evidence of initial multiple sparks is obvious at $T \approx 0.00683013$. However, the influence apparently becomes fuzzy at $T \approx 0.01366026$ and fuzzier at $T \approx 0.48121712$, while the maximal temperature,

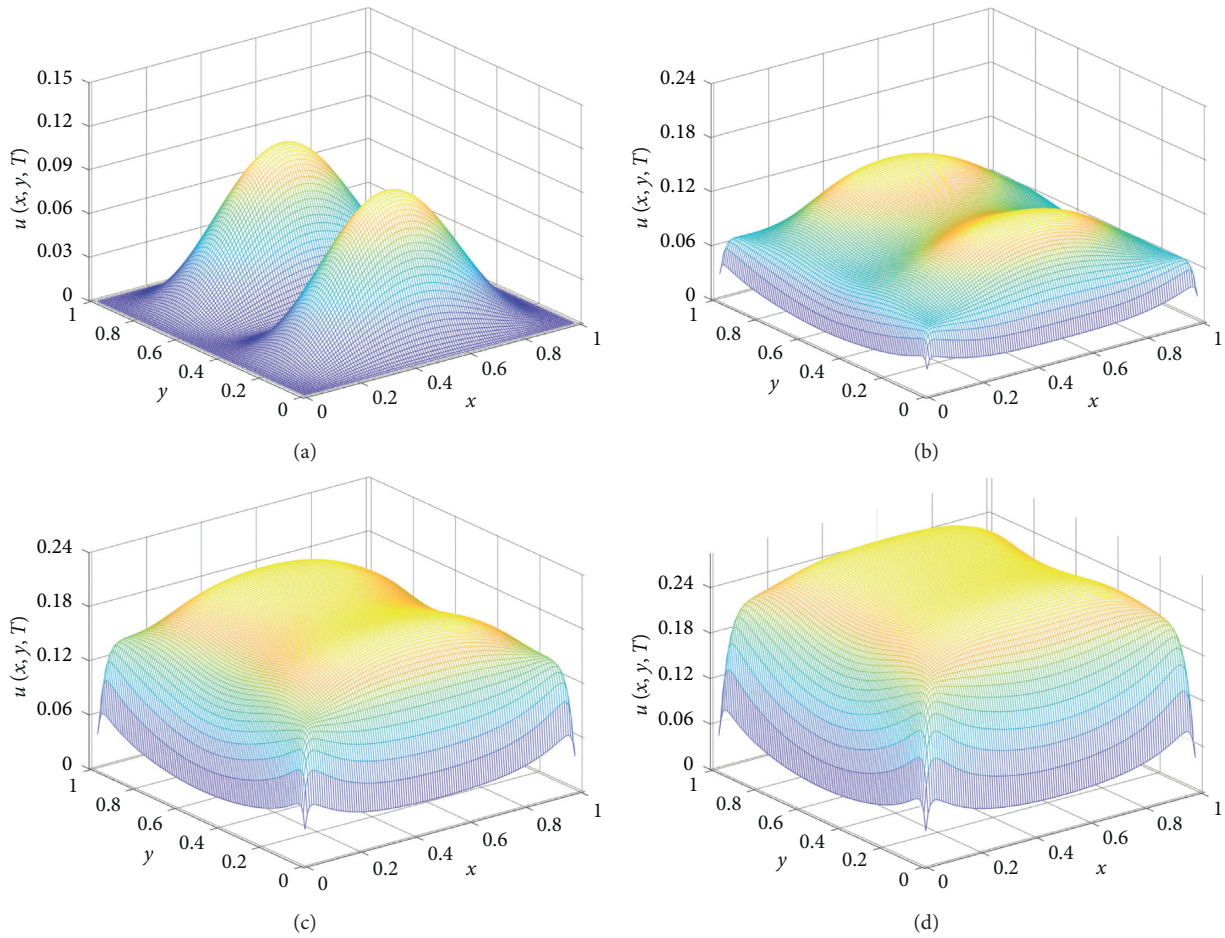


Figure 1: Numerical solution $u(x, y, T)$ at $T = 2\tau \approx 1.36602691 \times 10^{-5}$ (a); $T = 10,000\tau \approx 0.06830134$ (b); $T = 20,000\tau \approx 0.13660269$ (c); and $T = 30,000\tau \approx 0.20490403$ (d). The corresponding maximal values are 0.09996733, 0.13774652, 0.20421692, and 0.28487089, respectively.

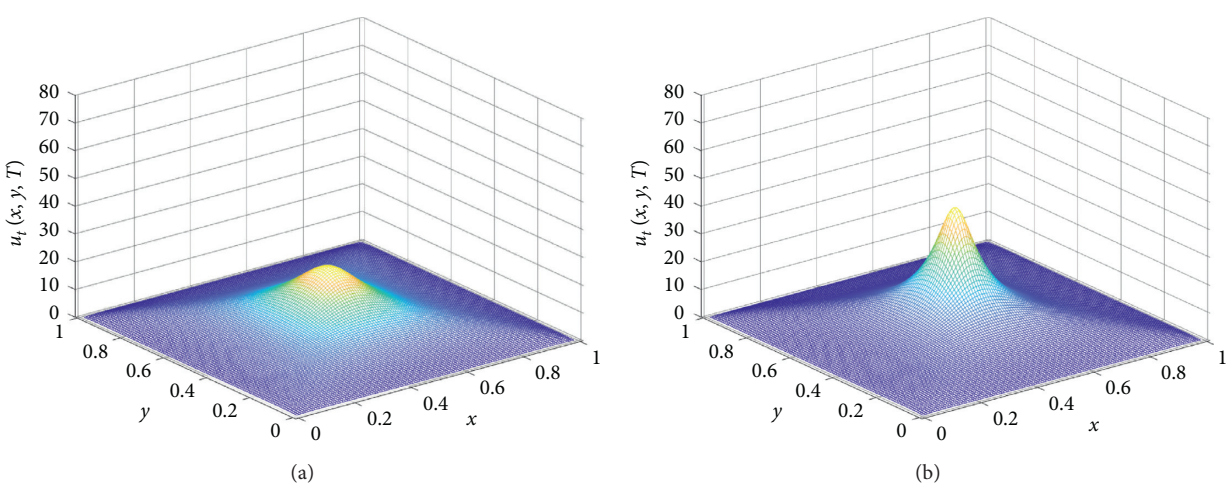


Figure 2: Continued.

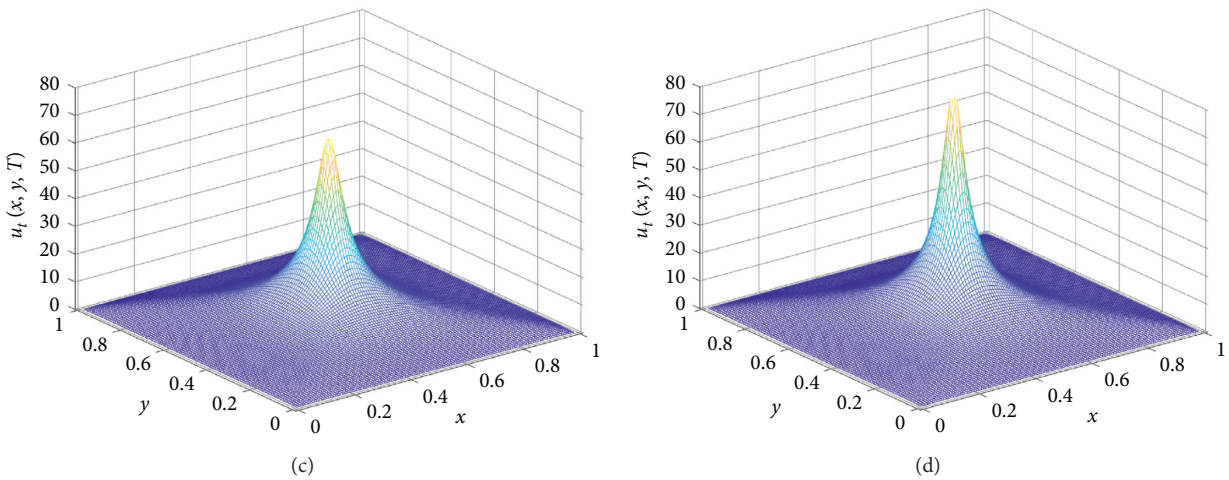


Figure 2: Temporal derivative function, or rate of temperature change, $u_t(x, y, T)$ prior to quenching at $T = 69,600\tau \approx 0.47537736$ (a); $T = 69,700\tau \approx 0.47606037$ (b); $T = 69,720\tau \approx 0.47619698$ (c); and $T = 69,726\tau \approx 0.47623796$ (d). The corresponding maximal values are 22.39748639, 43.58288207, 65.42874536, and 82.82435985, respectively.

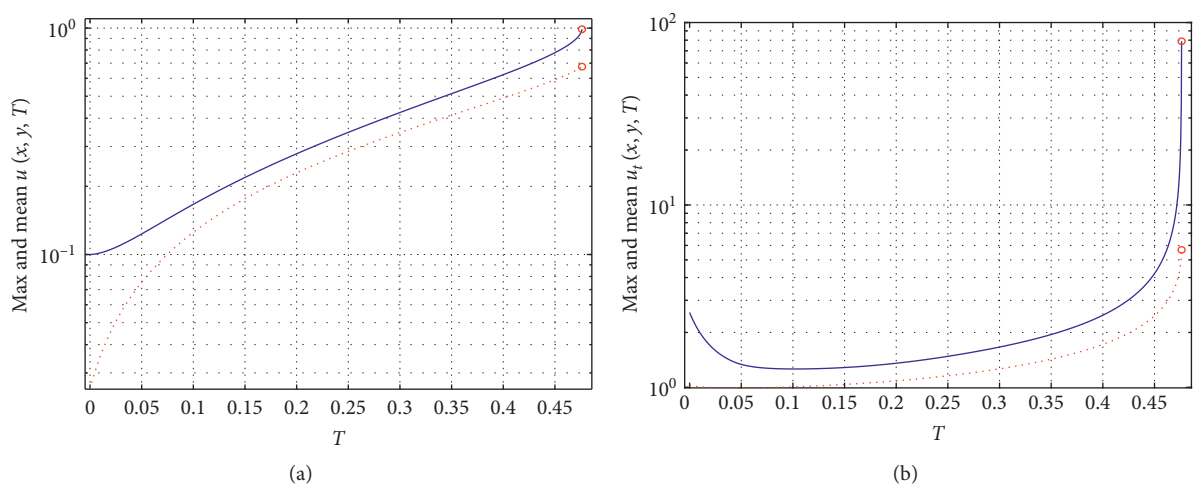


Figure 3: Trajectories of the maximal (blue curves), mean values (red dotted curves) of $u(x, y, T)$, $0 \le x, y \le 1$, (a); and the maximal (blue curves), mean values (red dotted curves) of $u_t(x, y, T)$, $0 \le x, y \le 1$, (b) for $T \in (0, 69725\tau]$. Explosive increments of the maximal derivative function values as t approaches the quenching time $68725\tau \approx 0.47623113 < T^* \approx 69727\tau \approx 0.47624479$ are again visible. This is consistent with quenching criteria (6) and (7).

$\max_{0\le x,y\le 1} u(x, y, T)$, increases monotonically from 0.13965273 to 0.20441452 and then 0.97828162. We also notice that the maximal value locations move from the initial four sparks to the central of the spatial domain. The phenomenon can be seen more clearly in the last figure of u, for which $T \approx 0.48142203$. This is further supported by the corresponding set of u_t surfaces in Figure 5. The facts indicate a unique quenching/blow-up position in the combustion chamber which is physically correct [4, 5, 11].

In the final phase at $T \approx 0.48142203$, values of $u \approx 1$, which is the singular point of the source function $f(u)$, in almost the entire combustion domain. On the other hand, we also observe that, while the distribution profile of u_t is similar to that of u, the amplitude of rate function u_t jumps to approximately 1.5×10^5. This is an indication that, during the final ignition, the fuel combustion speed must be tremendously high, while the dimensionless temperature stays under the unity.

We show both the maximal and mean values of the numerical solution u and its temporal derivative u_t, $(x, y) \in \overline{\mathcal{D}}$, in Figure 6. It is interesting that, while both mean values increase monotonically, the maximal derivative function decreases slightly in the beginning but quickly picks up the momentum and increases rapidly and becomes unbounded as $t \longrightarrow T^*$. However, both maximal values remain positive throughout computations. The simulation features well agree with the monotone solution property predicted in the theory of thermal combustion [2, 3].

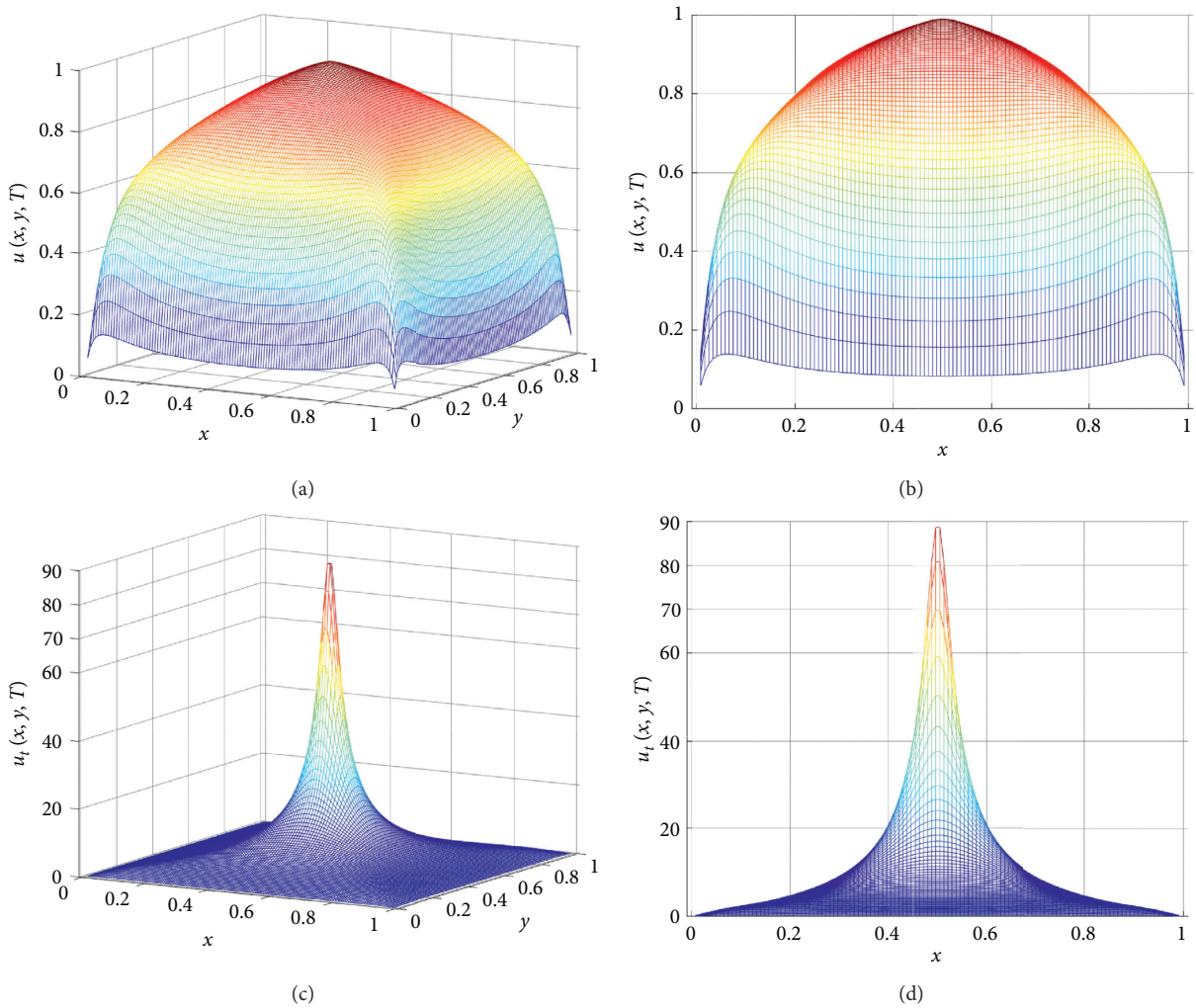


Figure 4: Numerical solution $u(x, y, T)$ (a, b) and its temporal rate of change function $u_t(x, t)$ (c, d), together with their respective $x - u/u_t$ plane projections, are given at the reference $T \approx 69727\tau \approx 0.47624479$ immediately before quenching. We have $\max_{0\le x,y\le 1} u(x, y, T) \approx 98952207$, $\max_{0\le x,y\le 1} u_t(x, y, T) \approx 87.32452308$. Again, plots are over internal mesh points without showing the boundary data (2).

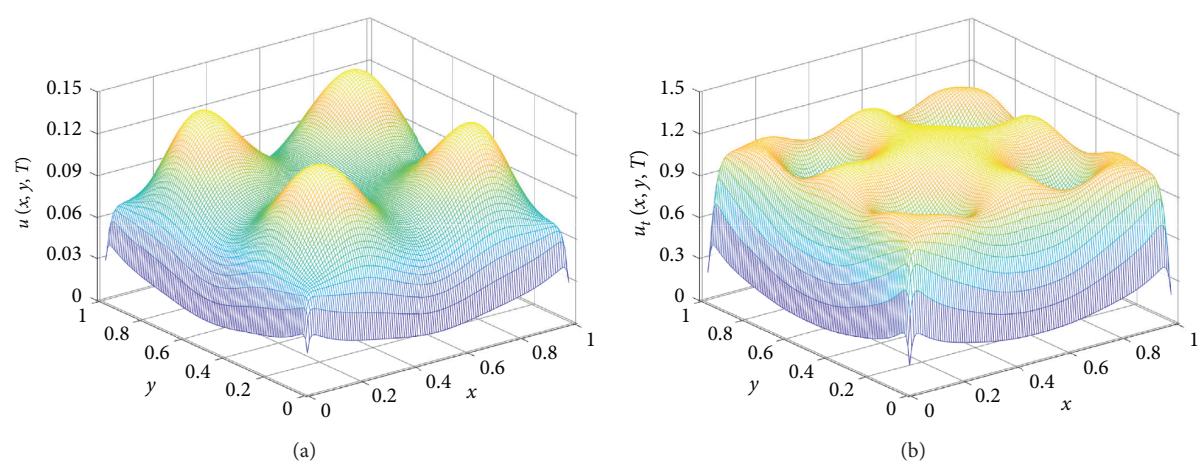


Figure 5: Continued.

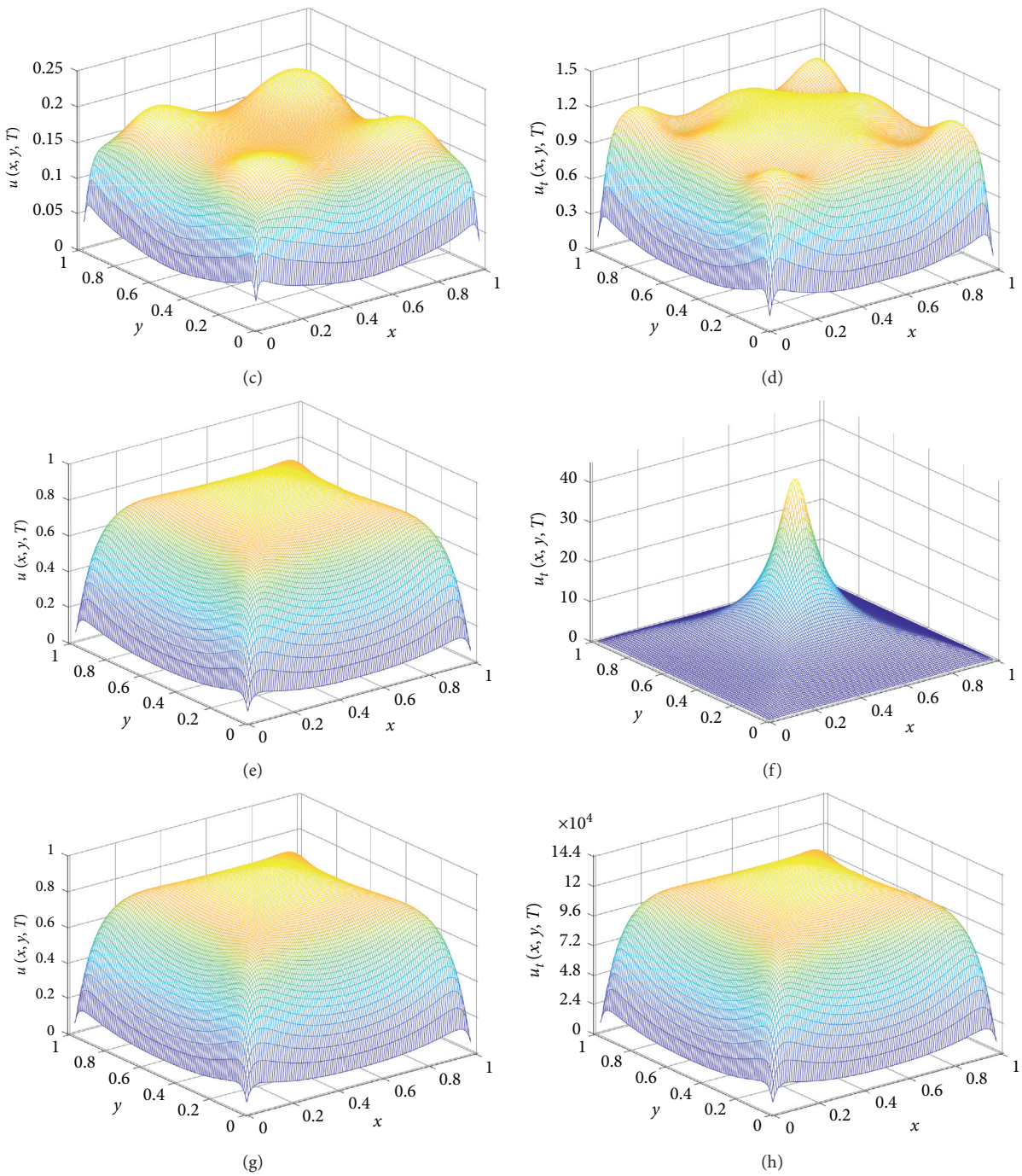


Figure 5: Numerical solution $u(x, y, T)$ (a, c, e, g) and its temporal rate of change function $u_t(x, t)$ (b, d, f, h) at references $T = 1000\tau \approx 0.00683013$, $2000\tau \approx 0.01366026$, $70455\tau \approx 0.48121712$, and $70485\tau \approx 0.48142203$, respectively. For the last pair of figures, we have $\max_{0\le x,y\le 1} u(x, y, T) \approx 0.99014166$, $\max_{0\le x,y\le 1} u_t(x, y, T) \approx 1.44966641 \times 10^5$. Again, boundary data are omitted for better clarity.

Finally, in Figure 7, profiles of the solution u and its temporal derivative u_t are given at $y = 0.5, T = 70483\tau \approx 0.48140837$ immediately before quenching. It can be seen that while the temperature inside the chamber increases uniformly towards the unity, the rate-of-change function u_t, in other words, the velocity of the temperature field evolution, grows rapidly to the infinite at the quenching location $(x^*, y^*) = (0.5, 0.5) \in \mathscr{D}$.

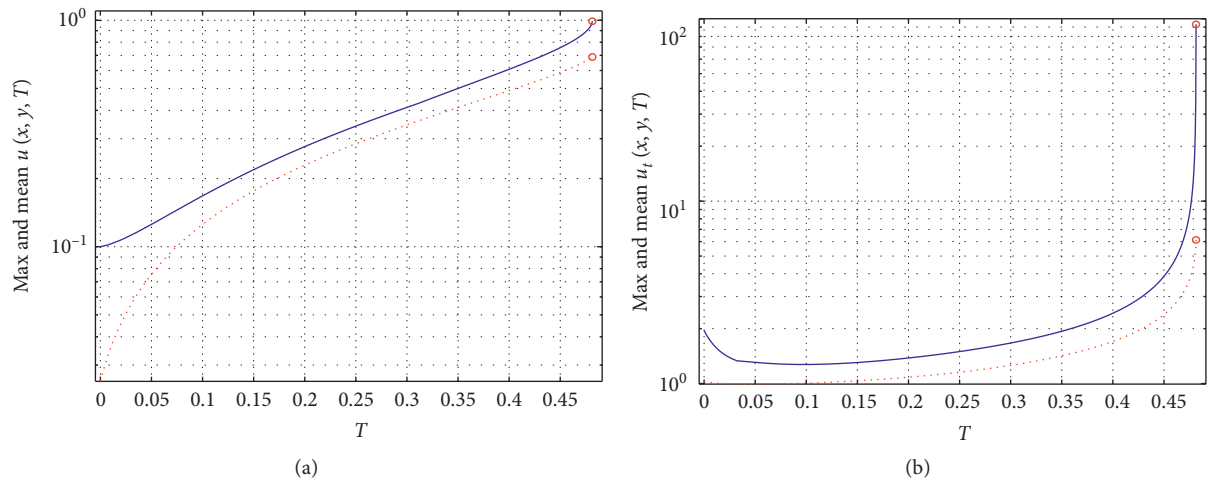


Figure 6: Trajectories of the maximal (blue curves), mean values (red dotted curves) of $u(x, y, T)$, $0 \le x, y \le 1$, (a); and the maximal (blue curves), mean values (red dotted curves) of $u_t(x, y, T)$, $0 \le x, y \le 1$, (b) for $T \in (0, 0.48140837]$. They are similar to those in Figure 3. Explosive increments of the derivative function values as t approaches the quenching time $70483\tau \approx 0.48140837 < T^* \approx 70485\tau \approx 0.48142203$ are again visible. The phenomena agree with (6) and (7) satisfactorily.

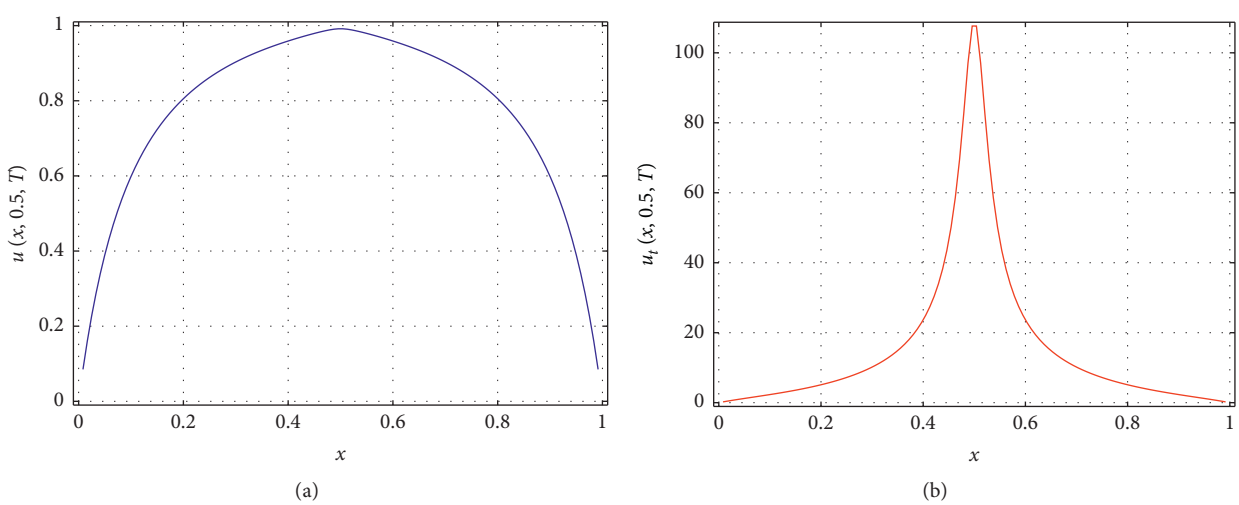


Figure 7: Profiles of the numerical solution $u(x, y, T)$ (a) and $u_t(x, y, T)$ (b), at $y = 0.5$, $T = 0.48141520$ before quenching. The profiles are highly symmetric with peaks at the center of the scaled dimensionless combustion domain. Boundary values are not included.

5. Conclusions and Future Endeavours

In this article, we have built and investigated two interconnected second-order finite difference schemes for the numerical solution of the stochastic quenching-combustion partial differential equation (1) equipped with Dirichlet boundary conditions and initial condition (2) and (3). This is for the first time, to the authors' best knowledge, to use a second-order finite difference scheme for effective approximations of quenching solutions on arbitrary spatial meshes. Numerical results obtained are not only consistent with existing results, but also offering more structure related details. Parameter determination procedures are discussed and analyzed.

Computer simulation experiments are focused on two examples with solid combustion-reaction backgrounds [1, 5]. Evenly distributed multiple sparks are deployed in combustion chambers. It is found that both the temperature and its temporal derivative distributions increase monotonically inside the chamber once the ignition starts. The evolution of the temperature field remains smooth and bounded during computations. However, on the other hand, the derivative distribution in the chamber increases rapidly and becomes unbounded as t reaches the neighborhood of the quenching time T^*. This is an indication of a physical combustion reaction of the fuel utilized.

The new second-order methods developed on arbitrary spatial grids are extremely straightforward and easy to use in programming implementations. They capture successfully the unique quenching location of the solution of (1)–(3). The numerical results acquired correctly reflect the fact that the final quenching-combustion location is unique and in the

center of the camber domain, rather than the multiple initial ignition spark positions suggested by initial functions.

Future endeavors in this direction involve continuing numerical analysis and the implementation of numerical algorithms for more rigorous stochastic structure coalitions and conditions. The authors will also pay special attention to the degeneracies, since these are indications of engine metal fatigue [12, 14]. They are extremely important in applications. Higher-order continuous and discontinuous Galerkin methods [15, 16] may also be explored for solving singular quenching problems including (1)–(3). It is also the authors' intention to inspire more collaborations in this very meaningful and promising territory of advanced theory and numerical methods for the natural world.

Data Availability

The data are available from the corresponding author upon any reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The first, second, and last authors acknowledge the constant support from Baylor University during the realization of this work. The third author acknowledges the financial support from the National Council for Science and Technology of Mexico (CONACYT) through grant A1-S-45928.

References

[1] M. C. Branch, M. E. Beckstead, T. A. Litzinger et al., "Nonsteady combustion mechanisms of advanced solid propellants," *Annual Technical Report, 94-05*, Center for Combustion and Environmental Research, University of Colorado, Boulder, CO, USA, 1994.

[2] V. I. Naoumov, "Part II: mathematical modeling of selected typical modes of combustion," *Chemical Kinetics in Combustion and Reactive Flows*, Cambridge University Press, London, UK, 2019.

[3] T. Poinsot and D. Veynante, *Theoretical and Numerical Combustion*, Edwards Publisher, Philadelphia, PA, USA, 2005.

[4] X. Wang, K. Tsuchiya, S. Fujita, S. Muto, and Y. Iijima, "Experiment and numerical simulation on quench characteristics of ReBCO-impregnated coil," *IEEE Transactions on Applied Superconductivity*, vol. 27, pp. 4–4700105, 2017.

[5] H. A. Levine, "Quenching, nonquenching, and beyond quenching for solution of some parabolic equations," *Annali di Matematica Pura ed Applicata*, vol. 155, no. 1, pp. 243–260, 1989.

[6] Q. Sheng and A. Q. M. Khaliq, *Linearly Implicit Adaptive Schemes for Singular Reaction-Diffusion Equations, Chapter 9, Adaptive Method of Lines*, Capman & Hall/CRC, London, UK, 2001.

[7] Q. Sheng, "Adaptive decomposition finite difference methods for solving singular problems-a review," *Frontiers of Mathematics in China*, vol. 4, no. 4, pp. 599–626, 2009.

[8] J. L. Padgett and Q. Sheng, "Convergence of an operator splitting scheme for abstract stochastic evolution equations," *Advances in Mechanics and Mathematics*, vol. 18, pp. 163–179, 2019.

[9] J. K. Hale, *Asymptotic Behavior of Dissipative Systems*, American Math Society, Philadelphia, PA, USA, 1988.

[10] H. Kawarada, "On solutions of initial-boundary value problems for," *Publications of the Research Institute for Mathematical Sciences*, vol. 10, pp. 729–736, 1975.

[11] H. A. Levine and J. T. Montgomery, "The quenching of solutions of some nonlinear parabolic equations," *SIAM Journal on Mathematical Analysis*, vol. 11, no. 5, pp. 842–847, 1980.

[12] M. A. Beauregard and Q. Sheng, "Solving degenerate quenching-combustion equations by an adaptive splitting method on evolving grids," *Computers & Structures*, vol. 122, pp. 33–43, 2013.

[13] M. A. Beauregard and Q. Sheng, "A fully adaptive approximation for quenching-type reaction-diffusion equations over circular domains," *Numerical Methods for Partial Differential Equations*, vol. 30, no. 2, pp. 472–489, 2014.

[14] J. L. Padgett and Q. Sheng, "Numerical solution of degenerate stochastic Kawarada equations via a semi-discretized approach," *Applied Mathematics and Computation*, vol. 325, pp. 210–226, 2018.

[15] R. M. Hafez and M. A. Zaky, "High-order continuous Galerkin methods for multi-dimensional advection-reaction-diffusion problems," *Engineering with Computers*, vol. 36, no. 4, pp. 1813–1829, 2020.

[16] M. A. Zaky, A. S. Hendy, and J. E. Macías-Díaz, "Semi-implicit Galerkin-Legendre spectral schemes for nonlinear time-space fractional diffusion-reaction equations with smooth and nonsmooth solutions," *Journal of Scientific Computing*, vol. 82, pp. 1–27, 2020.

[17] L. Zhu and Q. Sheng, "A note on the adaptive numerical solution of a Riemann-Liouville space-fractional Kawarada problem," *Journal of Computational and Applied Mathematics*, vol. 374, 2020.

[18] Q. Sheng, *The ADI Method, Chapter of Encyclopedia of Applied and Computational Mathematics*, Springer Verlag GmbH, Heidelberg, Germany, 2015.

[19] G. I. Marchuk, *Methods of Numerical Mathematics*, Springer-Velag, New York, NY, USA, 1982.

[20] A. Iserles, *A First Course in the Numerical Analysis of Differential Equations*, Cambridge University Press, Cambridge, NY, USA, 2nd edition, 2010.

[21] J. Kabre and Q. Sheng, "A preservative splitting approximation of the solution of a variable coefficient quenching problem," *Computers and Mathematics with Applications*, 2021, In press.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 5549188, 18 pages
https://doi.org/10.1155/2021/5549188

Research Article

Evacuation Model of Emotional Contagion Crowd Based on Cellular Automata

Qian Xiao [1] **and Jiayang Li** [1,2]

[1] *Department of Information Engineering, Shenyang University, Shenyang 110044, China*
[2] *Department of Business Administration, Northeastern University, Shenyang 110000, China*

Correspondence should be addressed to Qian Xiao; xiaoqian_neu@163.com

Received 30 January 2021; Revised 7 March 2021; Accepted 13 March 2021; Published 27 March 2021

Academic Editor: Jorge E. Macias-Diaz

Crowd evacuation under emergency is an important task of world public security research and practice. In order to describe the microemotional contagion of evacuation individuals, a cellular automata-based evacuation model of emotional contagion crowd based on the classical SIS model of infectious diseases is proposed in this paper. Firstly, the state of evacuation individual is defined as "emotional susceptible" and "emotional infective." Then, a dynamic model considering emotional contagion is established with cellular automata. Based on the models of static floor field and dynamic floor field, the emotion updating rules and state updating rules are constructed. The influence of perception domain radius on pedestrian evacuation process is analyzed through experiments. The conclusion can provide evacuation guidance for evacuation individuals. The comparative experiment results show that the improved model can reflect the movement characteristics of evacuation individuals effectively. The evacuation efficiency of the whole system is also effectively improved due to the consideration of emotional contagion and evacuation strategy.

1. Introduction

With the development of social progress and frequent economic, political, and cultural exchanges, people's activities are increasingly rich. In recent years, various accidents caused by the crowd have happened frequently. Therefore, how to grasp the characteristics of crowd evacuation behavior in emergency is the key to formulate effective guidance plan. In the past few decades, many researchers have mainly used experimental and simulation methods to study crowd evacuation behavior. The experimental methods mainly include the questionnaire method after the real event and controlled evacuation experiment with personnel participation. Afterwards, questionnaire method has information lag. Controllable experiments are often costly, and it is difficult to simulate dangerous scenarios, such as fire and terrorist attacks. Therefore, building a crowd evacuation simulation model in line with the actual situation has gradually become an important method to analyze the characteristics of crowd evacuation behavior, which has been widely used [1–3]. Simulation model method is to design reasonable building facilities and environment and reshape the basic attributes of the crowd. Then, take the crowd or individual as the research object to carry out the simulation of normal or emergency evacuation process. The evacuation simulation model can be used to evaluate the effectiveness of guidance measures, help the site management to make emergency plans, and provide theoretical support for emergency decision-making.

At present, the research of indoor crowd evacuation simulation model is mainly divided into macromodel and micromodel. The macrosimulation model usually adopts the classical graph theory method. The evacuation decision-making problem is transformed into an evacuation route selection problem. It usually adopts the minimum cost maximum flow method, the queuing theory method, and the shortest path method [4–7]. Micromodels usually take individuals as objects. It can identify many typical characteristics of individual movement or human evacuation behavior. Common models include cellular automata model,

social force model, lattice gas model, and visual obstacle model [8–13]. Compared with other models, cellular automata (CA) model has powerful complex computing function and highly dynamic characteristics. It has a strong ability to simulate the spatiotemporal dynamic evolution of space complex system and requires relatively less computation. Therefore, it has been widely used in crowd evacuation research.

Cellular automata were proposed by Blue et al. [14] in 1997, and then it began to be applied in pedestrian evacuation. The basic CA model represents the environment site as a grid composed of cells. The evacuation individuals move from one cell to the adjacent cell according to certain rules. Liu et al. [15] established an improved CA evacuation model considering the influence of the density of people around the exit on the evacuation behavior. Li et al. [16] proposed an extended CA model, considering the effect of sudden hazard. Zhang et al. [17] established a multiagent decision-making framework based on CA model to study the evacuation problem of stadium with or without obstacles. Yang et al. [18] present a new static floor field method based on cellular automaton for simulations of evacuation processes. Zheng et al. [19] proposed an extended CA model to study the effect of three-dimensional smoke diffusion on crowd evacuation. Geng et al. [20] proposed an extended cellular automata model to simulate the individual evacuation information under adverse line of sight conditions with uncertainty.

In emergencies, individual emotion plays an important role in evacuation decision-making. The emotion of the evacuation individual is dynamic and can be perceived and spread. Franovetter [21] proposed the concept of "threshold theory" and proposed that every event has an emotional threshold. The OCEAN model [22] proposed by McCrae et al. is the most widely used personality model, and the model describes individual personality from five vector dimensions. Helbing [23] simulated the dynamic escape behavior of evacuating individuals and concluded that escape behavior was a behavioral manifestation of emotion and confirmed that emotion among many people was contagious. Wang et al. [24] investigated the correlations between evacuation crowd characteristics and the evacuation emotion in subway emergencies by questionnaire. Shi et al. [25] proposed a new deterministic model of impatience, which described the impatience during pedestrian evacuation qualitatively and quantitatively. Wang et al. [26] considered that emotional contagion has an impact on the behavior of both individuals and groups and formulated the control problem of emotional contagion as an optimal control problem of crowd evacuation.

The human study of group emotion communication originated from infectious diseases. There are two kinds of scientific epidemic models to explore how diseases spread among groups, such as SIR model and SIS model [27, 28]. SIR model is to study the transmission relationship among "susceptible," "infective," and "removal." Fu et al. [29] modified the epidemiological SIR model and proposed a new (CA-SIRS) emotional contagion model. The SIS model considers that it is difficult for individuals to achieve complete immunity when considering the infection of some diseases. Therefore, there are only "susceptible persons" and "infected persons" in the process of infection. The SIS model describes that an individual can change from a susceptible state to an infected state. After cure, he is still in a susceptible state and has the probability of being infected. However, the total number of infected people in SIS model should remain unchanged, and it is not suitable to describe the micro-emotional contagion of pedestrian evacuation individuals. Therefore, this paper combines SIS model with CA model to build a crowd evacuation model, considering emotional contagion under emergency. Then, the evacuation process is simulated.

2. Basic Theory

2.1. Cellular Automata Model. Cellular automata are a kind of dynamic system with discrete time and space. It uses discrete space layout and discrete time interval to divide the cell into finite states. The evolution of the individual state of the cell is only related to its current state and the state of a local neighborhood. Each cell scattered in the regular grid takes a finite discrete state, follows the same action rules, and updates synchronously according to the determined local rules. The basic elements of cellular automata are as follows:

(1) Space: the spatial lattice of cells in space can be one-dimensional, two-dimensional, or multidimensional.

(2) State set: there are only two different states, which can be coded as 0 and 1, respectively, or represented by different colors.

(3) Neighbor: there is a cell with radius of 1 around a cell, which can affect the state of the cell at the next moment. The most common neighbors are Moore type and von Neumann type.

(4) According to the state of the cell and its eight neighboring cells, the dynamic function or state transfer equation of the cell state at the next moment is determined.

2.2. Typical Models of Infectious Diseases. Emotional contagion can be studied with reference to the disease transmission model. There are two kinds of traditional infectious disease models: SIR model and SIS model.

SIR model refers to three types of individual status: "susceptible," "infective," and "removal." S, I, and R, respectively, represent the proportion of the three groups in the total group. SIR model does not calculate the new birth rate and death rate in the process of infection, so $S + I + R = 1$. The model reflects the macroprocess of human infection: when infectious diseases occur, the total number of three groups of people is constant. Once "susceptible" (S) contacted with "infective" (I) will be infected with probability; "infective" (I) will change to "removal" (R) with a certain probability; "removal" is equivalent to vaccine injector and will no longer be infected with disease immunity.

The other classical SIS model considers that when some diseases are infected, it is difficult for individuals to achieve complete immunity. Therefore, there are only "susceptible"

and "infective" in the process of infection. SIS emotional contagion model describes that an individual changes from an easily infected state to an infected state. After being cured, he is still an easily infected individual and will enter the infectious system again. The differential system dynamics equation of the model is shown in equation (1). The state transition diagram is shown in Figure 1.

$$S + I = 1,$$
$$\frac{dS}{dt} = -\beta SI + \gamma I, \tag{1}$$
$$\frac{dI}{dt} = \beta SI - \gamma I,$$

where β is the daily contact infection rate; γ is the daily cure rate.

Since the SIS model adopts the macroscopic model of infection dynamics expressed by differential equation, it is not suitable for describing the microemotional contagion of evacuation individuals. In addition, the SIS model assumes that the total number of people does not change during the process of infection. However, when pedestrians evacuate, once they leave the exit, it is a safe area. That is, the number of people in the dangerous area at every moment is dynamic. Based on the above, this paper combines the classical SIS model with cellular automata to build a dynamic pedestrian evacuation model, considering emotional contagion in emergencies.

3. Evacuation Model of Emotional Contagion Crowd Based on Cellular Automata

3.1. Basic Cell Definition of Cellular Automata. In this paper, cellular automata model is used to simulate the evacuation situation in independent space. The space is L in length and W in width. It is divided into cell lattices of 0.4×0.4 for one person to stand. Each cell is represented by a set of number pairs, such as cell (x, y). The distance between adjacent cells is cl = 0.4m. The number of exits is m. The width is el_i, where

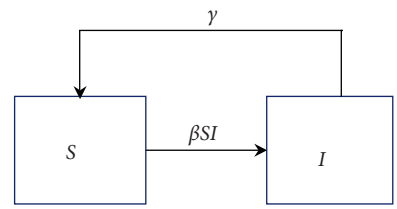


Figure 1: Schematic diagram of SIS model state transition.

$i \in [1, 2, \ldots, m]$. When the initial state $t = 0$, pedestrians are randomly distributed in the space, each occupying a cell. After the evacuation, Moore neighborhood rule with radius of 1 is adopted in the neighborhood of the cell, as shown in Figure 2. The time step is $\Delta t = 0.4s$. Each time, the pedestrians determine the next move strategy according to the transfer probability of the cell in each direction until they reach the exit and escape.

Burstedde et al. [30] introduced the concepts of static floor field and dynamic floor field in the CA model. Many scholars added friction, inertia, force, obstacle, and other factors to its expansion. The probability of pedestrian moving to its adjacent cell lattice is determined by static floor field, dynamic floor field, and their interaction. This paper analyzes the transition probability of static floor field and dynamic floor field, respectively. In this paper, the model is established on the basis of static field and dynamic field.

3.2. Static Floor Field. The static floor field (SFF) in cellular automata is the distance from each cell to the effective exit. Let $S_{ji}(t)$ denote the shortest distance from the neighborhood cell i of evacuation individual j to the exit at time t. Let the location coordinate of evacuation individual j be $C(x, y)$, and the location coordinate of neighborhood cell i be $C(x_i, y_i)$. There are m effective exits in the evacuation space. The midpoint coordinate of each exit is $C(x_{e_k}, y_{e_k})$. According to the static floor field of cellular automata, the transition probability formula is shown in the following equation:

$$S_{ji} = \max_{i \in \text{all neighbors}} \left\{ \min_{i \longrightarrow \text{all exit } k \in [1,m]} \left\{ \left[\left(x_{e_k} - x_i\right)^2 + \left(y_{e_k} - y_i\right)^2 \right]^{1/2} \right\} \right\} - \min_{i \longrightarrow \text{all exit } k \in [1,m]} \left\{ \left[\left(x_{e_k} - x_i\right)^2 + \left(y_{e_k} - y_i\right)^2 \right]^{1/2} \right\}, \tag{2}$$

where the second term of the formula represents the minimum value of the neighborhood cell distance among all exits. The first term represents the maximum value of the minimum value of the neighborhood cell distance among all exits. The formula is a two-level progressive relation. The first layer is to calculate the shortest distance between eight cells in the neighborhood and all effective exits. The second layer is to compare the deviation between the shortest distance of eight cells in the neighborhood and the maximum of the shortest distance. If it is equal to the maximum value of the shortest distance, then $S_{ji}(t) = 0$. If the deviation from the maximum value of the shortest distance is greater, then the cell has the shortest distance from the exit. That is, the greater the $S_{ji}(t)$ value is, the greater the probability of the cell being selected.

3.3. Dynamic Floor Field. In the emergency situation, the state of the evacuation individual changes after the emotion reaches a certain threshold. Thus, it causes the change of behavior strategy. The behavior execution process of evacuation individual j at time t is shown in Figure 3. Therefore, the dynamic field of this model mainly includes emotion updating rules and state updating rules.

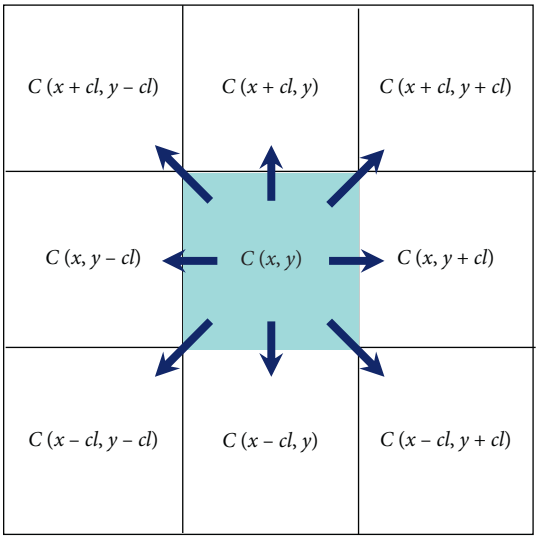


Figure 2: Moore neighborhood rule and direction transition probability.

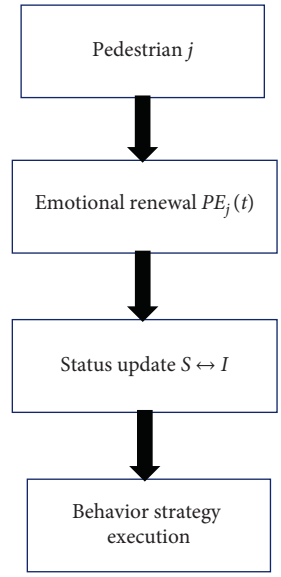


Figure 3: Diagram of an iterative process of evacuation individual j at time t.

3.3.1. Emotional Contagion Function. The emotional intensity of evacuation individual j fluctuates with time. The dynamic equation is shown in the following equation:

$$\mathrm{PE}_j(t) = \mathrm{PE}_j(t-1) + (1-\xi)\cdot \Delta e_j(t) + \tau, \tag{3}$$

where $\mathrm{PE}_j(t-1)$ is the emotional intensity of evacuation individual j at time $t-1$; $\Delta e_j(t)$ is the emotion increment of evacuation individual j at time t; ξ is the perceptual attenuation coefficient; and the value range is $[0,1]$. It indicates that there is a certain attenuation in the superposition of emotion increment of evacuation individuals. τ is the random emotional disturbance of individual j at time t.

\Equation (3) shows the renewal mechanism of emotion and the evolution mechanism of infection. The emotional intensity $\mathrm{PE}_j(t)$ is based on the self-emotional intensity at time $t-1$ and superimposed with the emotional influence of others in the surrounding perception area.

3.3.2. Emotional Perception Increment. The emotional fluctuation of evacuation individuals is related to their own speed and the emotion of others around them. Therefore, this model defines the emotional increment function of evacuation individual j at time t, as shown in the following equation:

$$\Delta e_j(t) = \Delta f_j(v,t) + \Delta\varphi_j(t), \tag{4}$$

where $\Delta f_j(v,t)$ is the internal emotional change of evacuation individual j caused by speed at time t; $\Delta\varphi_j(t)$ is the emotional change of others perceived by individual j at time t.

(1) Definition of function $\Delta f_j(v,t)$: in the process of evacuation, if the evacuation speed is slower than the maximum expected speed, the negative emotions (irritability, tension, and anxiety) increase faster. If the walking speed of the evacuation individual is close to the maximum expected speed, the growth of their bad mood is relatively slow [24]. Therefore, the emotional increment function caused by speed is shown in the following equation:

$$\Delta f_j(v,t) = \left(1 - \frac{v_j(t-1)}{v_j^{\max}}\right)\cdot \exp\left(-\alpha\cdot\left(\frac{v_j(t-1)}{v_j^{\max}}\right)\right), \tag{5}$$

where $v_j(t-1)$ represents the speed of the previous movement of the evacuation individual j at time t; $v_j^{\max}$ is the maximum expected speed of individual j during evacuation; α is the velocity sensitive parameter of evacuation individual j at time t; and the value range is $[0,1]$.

This function can clearly express the influence of the difference between the previous moving speed and the maximum expected speed on the emotional fluctuation. If the moving speed is close to the maximum expected speed, it will cause less fluctuation of individual emotion. On the contrary, if the moving speed is far less than the maximum expected speed, the $\Delta f_j(v,t)$ value will be close to 1. In other words, the degree of emotion will increase.

(2) The definition of function $\Delta\varphi_j(t)$: Zhang et al. [31] proposed that there were circular perception domains and perception domain radius for evacuation individuals. The traditional perception domain discretization thinks that the emotional state of each individual in the perception area can be fully perceived by the individual, but the actual situation is not so. Geng et al. [20] considered that the influence of information communication in line of sight is different from that in nonline of sight. At the same time, this model considers that the scope of emotional perception is not limited to contact perception. For individuals, the emotions of the surrounding people will be perceived, especially the emotions in the visual area will be more intense through visual perception. Therefore, this model divides the perception area into visual perception domain RV and non-visual perception domain RI. It is considered that the

emotion of others in the visual domain can be fully perceived, while the emotion in the nonvisual domain can be perceived through pushing, sound, action, etc., and the perception degree will decline to a certain extent.

Li et al. used angle $\pi/4$ to represent the visual range of pedestrians in eight neighborhood directions [32]. Usually the horizontal angle of people's binocular vision is about $3\pi/4$. $\pi/2$ is selected as the angle of visual area in this model, considering that people's field of vision will shrink under emotional fluctuations. Then, the parameters of the visual perception domain RV and the nonvisual perception domain RI are defined as follows:

① $\text{Cell}_j(x, y)$ is the location of evacuation cell j

② $\theta_j(t)$ is the angle radian between the forward direction $\text{dir}_j(t-1)$ of the evacuation individual cell j and the horizontal axis of the coordinate. Its value can be $(0, \pi/4, \pi/2, 3\pi/4, \pi, 5\pi/4, 6\pi/4, 7\pi/4)$

③ $\text{dir}_j(t-1)$ is the forward direction of evacuation individual cell j at the last moment. The corresponding relationship with $\theta_j(t)$ is shown in Figure 4

④ The visual area of evacuation individual is a sector with arc $\pi/2$

⑤ R is the emotional perception domain radius of evacuation individuals

The definition of visual perception domain RV: the sector with the cell coordinate $\text{cell}_j(x, y)$ of evacuation individual j at time t as the center, $[\theta_j(t) - \pi/4, \theta_j(t) + \pi/4]$ as the angle, and R as the radius is the visual perception domain RV of the evacuation individual at time t. As shown in Figure 5, the green color area is the visual perception domain after discretization. The emotions of the evacuated individuals in this area can be perceived.

Definition of nonvisual perception domain RI: the circle with the cell coordinate $\text{cell}_j(x, y)$ of evacuation individual j at time t as the center, R as the radius, and removing the part of RV is the nonvisual perception domain RI of the evacuation individual at time t. As shown in Figure 5, the gray area is the nonvisual perception domain. The emotion of the evacuation individual is partially perceived.

Therefore, the visual perception domain and nonvisual perception domain, corresponding to different forward directions of evacuation individual cell j at time t, are shown in Figure 6.

According to the perception domain, the evacuation space is divided into three situations: exit perception area A_1, adjacent edge area A_2, and other area A_3, as shown in Figure 7.

As shown in Figure 7, exit perception area A_1 is composed of two squares with exit endpoint as vertex and R as side length and a rectangle with length R and width el. When evacuation individuals are in this area, they are not affected by other people's emotions, because the exit is the most attractive. Therefore, it can be considered that when in the exit perception area A_1, the emotion of the evacuation individual no longer fluctuates, such as the location of n_1. The perception area of evacuation individuals in the adjacent edge area A_2 is an incomplete circle, which needs to be

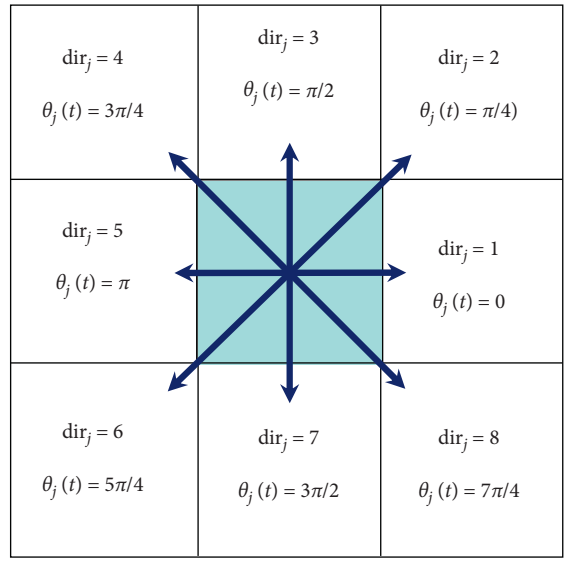


Figure 4: Corresponding relationship between forward direction and included angle radian.

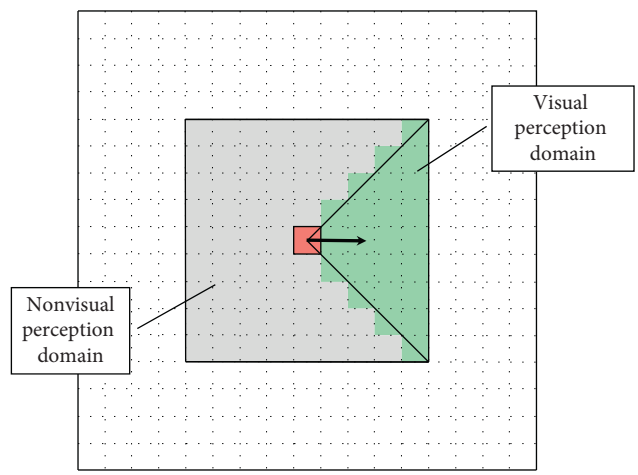


Figure 5: Visual perception domain and nonvisual perception domain when $R = 2$m.

simulated according to the incomplete rules, such as the location of n_2 and n_3. When the evacuation individual is in the other area A_3, the visual perception domain and nonvisual perception domain are complete, and the simulation calculation can be carried out according to the complete perception domain, such as the location of n_4 and n_5.

Nilsson et al. used empirical experiments to analyze other people's influence and behavior choice during indoor evacuation. They proved that the influence of close distance on evacuation individual's emotion and behavior is greater than that of long distance [33]. Fu et al. [29] also believed that the evacuation individuals were more sensitive to the emotional perception of nearby areas during the evacuation process. In this model, the influence of other people's emotion in visual perception domain and nonvisual perception domain on the evacuation individual is different. Combined with the distance factor, this model considers that

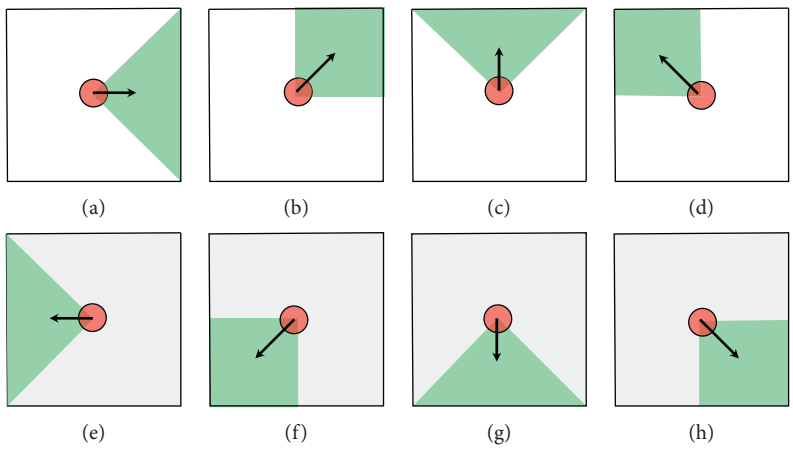


Figure 6: Possible emotion perception domain of evacuation individual j at time t.

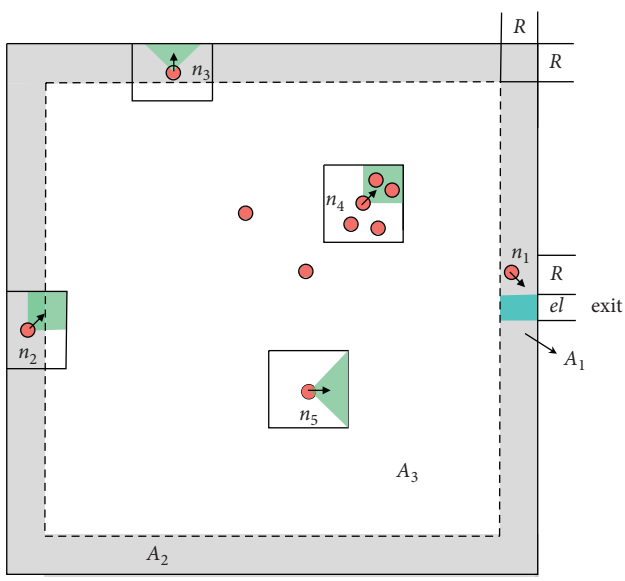


Figure 7: Area division of evacuation space with emotion perception.

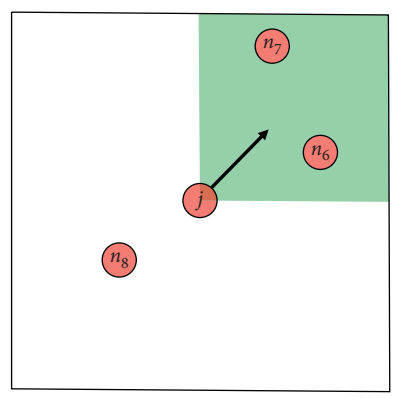


Figure 8: Emotional perception domain of individual j.

the perceived degree of others' emotion in the visual perception domain is greater than that in the visual perception domain and is greater than that in the nonvisual perception domain. Taking Figure 8 as an example, evacuation individual j is more likely to be affected by n_6, then by n_7, and finally by n_8.

In reference [29], the formula of distance as influence weight in emotion gain function is $[1-(1+\exp(-d_{jk}))^{-1}]$, and $d_{jk}=\sqrt{(x_j-x_k)^2+(y_j-y_k)^2}$ is the distance from evacuation individual j to target cell k. This formula reflects that evacuation individual j is more sensitive to the emotion of close evacuation individual, and the emotion of others far away from evacuation individual j has less influence. In order to reflect the difference of perception degree in three different situations, this model redefines the distance weight as the following equation:

$$dd_k=\left[1-\frac{1}{1+\exp\left(-d_{jk}\cdot cl\right)}\right]\bullet\left\{\sum\left[1-\frac{1}{1+\exp\left(-d_{jk}\cdot cl\right)}\right]\right\}^{-1}. \tag{6}$$

The average change function of other people's emotion perceived by evacuation individual cell j at time t is shown in the following equation:

$$\Delta\phi_j(t)=(g_1)^{-1}\bullet\sum_{k\in\mathrm{RV}}dd_k\times\left[\mathrm{PE}_k(t-1)-\mathrm{PE}_j(t-1)\right]+\eta\bullet(g_2)^{-1}\sum_{k\in\mathrm{RI}}\left[\mathrm{PE}_k(t-1)-\mathrm{PE}_j(t-1)\right], \tag{7}$$

where g_1 is the number of evacuees in the visual perception domain RV of evacuees cell j; g_2 is the number of evacuation individuals in the nonvisual perception domain RI of evacuation individual cell j; η is the degree to which the emotional difference of other people in the nonvisual domain is received by the evacuation individual j. Its value is

not greater than the minimum distance weight in the visual perception domain, that is, $0 \le \eta \le \min_{k\in \mathrm{RV}} dd_{jk}$.

The first term of equation (7) represents the average increment of emotional change of evacuation individual j in the visual domain at time t. It also shows that the amount of emotional change is related to distance, and the emotional difference in close distance has a great influence on evacuation individual j. The latter term indicates that the emotional change of evacuation individual j in the nonvisual domain is the overall weakening due to the absence of specific situation.

In conclusion, the emotional increment function of evacuation individual j at time t is updated as shown in the following equation:

$$\Delta e_j(t) = \Delta f_j(v,t) + \Delta\phi_j(t) = \left(1 - \frac{v_j(t-1)}{v_j^{\max}}\right) \bullet \exp\left(-\alpha \bullet \left(\frac{v_j(t-1)}{v_j^{\max}}\right)\right) + (g_1)^{-1} \bullet \sum_{k\in \mathrm{RV}} dd_k \times \left[\mathrm{PE}_k(t-1) - \mathrm{PE}_j(t-1)\right] + \eta \bullet (g_2)^{-1} \sum_{k\in \mathrm{RI}} \left[\mathrm{PE}_k(t-1) - \mathrm{PE}_j(t-1)\right]. \tag{8}$$

The incremental function of emotional intensity indicates that the change of speed will cause the emotional fluctuation of evacuation individuals. When the travel speed is close to the maximum expected speed, the fluctuation of emotion is small. The greater the difference between the travel speed and the maximum expected speed, the greater the emotional fluctuation. The emotions of others around also affect the emotional fluctuation of evacuation individuals [29]. The emotional influence of others in the visual perception domain is greater than that of people in the nonvisual perception domain. That is, the emotional influence of others in the visual perception domain is related to distance [24], and the emotion of evacuation individuals in close distance is easy to be perceived. It can be seen that the constructed incremental function of emotional intensity is more in line with the actual changes of individual emotions.

3.3.3. Mechanism of Emotional Renewal. The iterative process of evacuation strategy of evacuation individual j at time t is to judge the state update by emotion update and implement evacuation strategies corresponding to different states. The emotion of evacuation individual j at t time is related to the value of emotion, action strategy, state, and direction at time $t-1$. It is also affected by the emotional intensity of others within the range of emotion perception. The change of emotion will cause the dynamic change of evacuation individual's state and behavior strategy. The specific state update process is shown in Figure 9.

3.3.4. State Update Mechanism. Emotion gathered to a certain extent, resulting in the change of evacuation individual state. According to the SIS model, the evacuation state of the individual is defined as "emotional susceptible" and "emotional infective." The state transition rules follow SIS model. The initial definitions of "emotional susceptible" and "emotional infective" are as follows: when the emotional intensity $\mathrm{PE}_j(t)$ of individual j exceeds the threshold $\lambda_{\mathrm{emotion}}$ at time t, it is transformed into "emotional infective" by probability β. When the emotion fluctuation of "emotional infective" is lower than the threshold $\lambda_{\mathrm{emotion}}$, it is transformed into "emotional susceptible" by probability γ. In the process of evacuation, S is the proportion of "emotional susceptible" and I is the proportion of "emotional infective." Its dynamic differential dynamic equation and system conservation are the same as equation (1).

Different evacuation individuals will adopt different evacuation strategies because of their different states. Literature [34] analyzed the 2010 German "Love Parade" stampede accident video and observed that the crowd presented two different behavioral characteristics, namely, "positive characteristics" and "negative characteristics." People with "positive characteristics" will actively seek exit, accelerate forward, and have obvious herd behavior; people with "negative characteristics" are slow, try to avoid the crowd, and have no obvious herd behavior. Based on this, this model maps pedestrian emotional intensity to pedestrian behavior. During the evacuation, the "emotional susceptible" (S) adopts the "calm" strategy BS, the speed is vs, and tends to choose the direction with low density ρ; while the "emotional infective" (I) adopts the "impulsive" strategy BI, the speed is vi, and tends to choose the direction with high density ρ.

$\mathrm{PE}_j(t)$ is defined as the emotional intensity of evacuation individual j at time t, and the value range is $(0,1]$. That is, the emotion of evacuation individual will not disappear in the whole process of evacuation. According to the different emotional intensity, the evacuation behavior is different. $\lambda_{\mathrm{emotion}}$ is defined as the threshold value of emotional intensity change. It is the critical value from psychology to behavior, and the value range is $(0,1]$. The mapping relationship between emotional intensity and individual behavior type is shown in Table 1.

3.3.5. Transfer Strategy. In this model, cellular automata evacuation model is used. In each time step, the movement direction of the individual who does not leave the evacuation space is determined according to the next time movement probability. Let $P_{ji}(t)$ denote the probability of individual j transferring to cell C_i at time t. The evacuation individual will choose the neighborhood cell with the maximum transition probability as the target and move at the speed v_j. According to the theory of cellular automata, the transition probability is related to three basic characteristics such as the

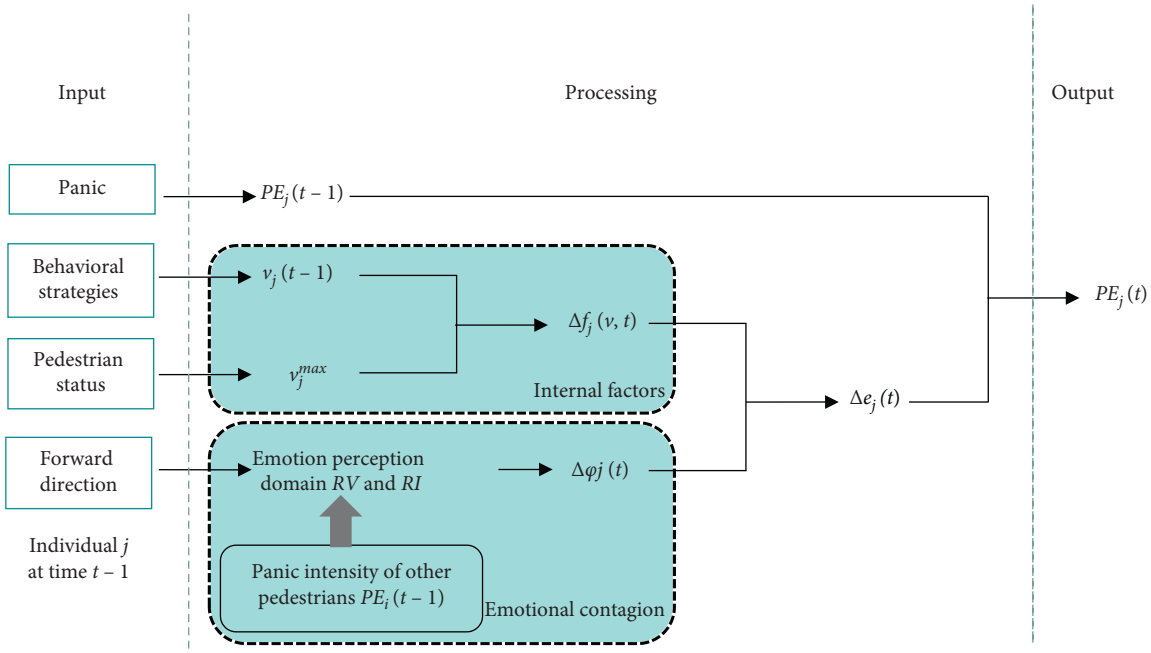


Figure 9: Emotional renewal mechanism of evacuation individual j at time t.

Table 1: Mapping relationship between emotional intensity and behavior types of evacuation individuals.

	$\mathrm{PE}_j(t) \le \lambda_{\mathrm{emotion}}$	$\mathrm{PE}_j(t) > \lambda_{\mathrm{emotion}}$
Types of individual evacuation status	S	I
Behavior types	"Calm"	"Impulsive"
Behavioral strategies	BS (vs, ρ)	BI (vi, ρ)

distance from neighborhood cell C_i to the exit, whether it is an obstacle cell. At the same time, this model considers that the crowd density of the forward direction is also a factor that determines the selection of target cells. Therefore, the cell transition probability is constructed as shown in the following equations:

$$P_{ji}(t) = U \bullet \exp\left(k_s S_{ji}(t) + k_\rho \rho_{ji}(t) \cdot \mathrm{cp}_{ji}(t) \cdot \mathrm{co}_{ji}(t)\right), \tag{9}$$

$$U = \left[\sum_{i \text{ is all neighbors}} \exp\left(k_s S_{ji}(t) + k_\rho \rho(t)\right) \cdot \mathrm{cp}_{ji}(t) \cdot \mathrm{co}_{ji}(t)\right]^{-1}, \tag{10}$$

where k_s is the influence coefficient of distance to the exit; k_ρ is the influence coefficient of density.

$\mathrm{cp}_{ji}(t)$ is the state parameter of neighborhood cell i of evacuation individual j at time t. If it has been occupied by evacuation individuals, the value is 0; otherwise, it is 1, as shown below:

$$\mathrm{cp}_{ji}(t) = \begin{cases} 0, & C(x_i, y_i) \text{ isn't empty} \\ 1, & \text{others}, \end{cases} \tag{11}$$

where $\mathrm{co}_{ji}(t)$ indicates whether the neighborhood cell i of evacuation individual. at time t is an obstacle. If it is an obstacle, the value is 0; otherwise, it is 1, as shown below:

$$\mathrm{co}_{ji}(t) = \begin{cases} 0, & C(x_i, y_i) \text{ is barrier}, \\ 1, & \text{others}, \end{cases} \tag{12}$$

where $\rho_{ji}(t)$ is the density around the neighborhood cell i of the evacuation individual j at time t. At present, all other evacuation individuals in the visible area of evacuation individuals will affect their choice. This model gives the density influence function of different types of groups combined with the decision types. Therefore, the transfer probability function of different types of crowd is different.

(1) "Emotional susceptible" S adopts "calm" strategy BS. The maximum expected speed is $v_j^{\max} = 1$ m/s. That is, when each step is updated, it moves 1 step. The target cell of "calm type" evacuation is 8 cells around, as shown in Figure 10.

At the same time, the density range of "calm type" evacuation individuals is the close density around the neighborhood cell i. As shown in Figure 10, the red cell is

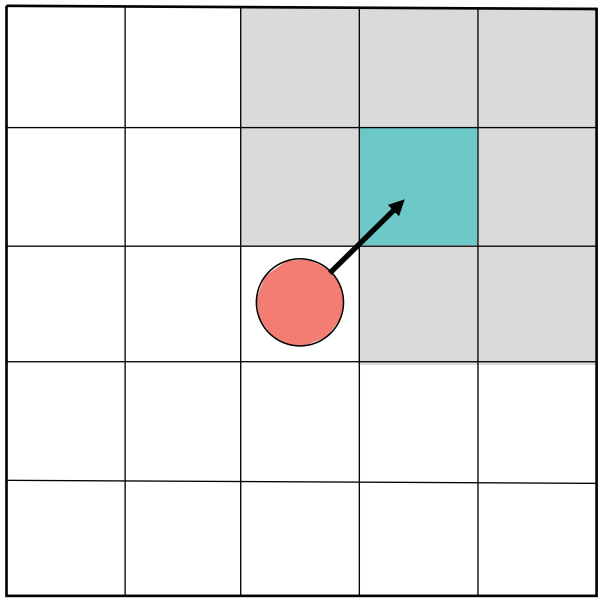


Figure 10: Density range observed for "calm."

evacuation individual j, and the blue cell is neighborhood cell i. When analyzing the density of i, evacuation individual j considers the density of the gray cell. Because the "calm" evacuation individuals described in this model choose the direction with low crowd density, the transfer probability formula of this kind of crowd is as follows:

$$P_{ji}(t) = U \bullet \exp\left(k_s S_{ji}(t) + k_\rho\left(1 - \rho_{ji}(t)\right) \cdot cp_{ji}(t) \cdot co_{ji}(t)\right)\left[\sum_{i \text{isallneighbors}} \exp\left(k_s S_{ji}(t) + k_\rho\left(1 - \rho(t)\right)\right) \cdot cp_{ji}(t) \cdot co_{ji}(t)\right]^{-1} \bullet \exp\left(k_s S_{ji}(t) + k_\rho\left(1 - \rho_{ji}(t)\right) \cdot cp_{ji}(t) \cdot co_{ji}(t)\right). \tag{13}$$

(2) "Emotional infective" I adopts "impulsive" strategy BI. The velocity is vi. They tend to choose the direction with high crowd density ρ. The maximum expected speed is $v_j^{\max} = 2\text{m/s}$. When each step is updated, it moves 2 steps. Then, the target cell of "impulsive" evacuation is two-step cell in each direction. Suppose the group of target cells is set G_1, as shown in the blue cells in Figure 11(a). At this time, the moving speed is $vi = v_j^{\max}$. If all the preferred eight cells are occupied, the target cell with the largest transition probability of eight cells in neighborhood is selected. Suppose the neighborhood cell is set G_2, as shown in the blue cells in Figure 11(b). At this time, the moving speed is $vi = 1\text{m/s}$.

The radius of density range of "impulsive" evacuation individuals is two steps, as shown in Figure 12.

Because the "impulsive" evacuation individuals described in this model choose the direction with high crowd density, the calculation method of the transition probability of this kind of crowd is as follows:

Calculate $P_{ji}(t)$ according to equation (9). If

$$\underset{i \in G_1}{P_{ji}(t)} \quad \sum_{i \in G_1} P_{ji}(t) = 0, \tag{14}$$

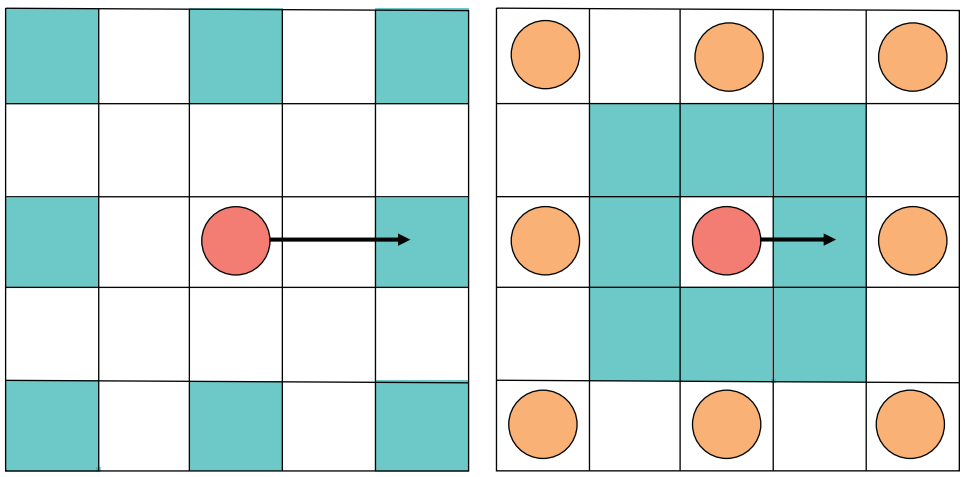
then calculate $\underset{i \in G_2}{P_{ji}(t)}$.

4. Simulation Experiment

4.1. Simulation Process. According to the emotional intensity of individual evacuation, the individual evacuation is divided into two groups: "emotional susceptible" and "emotional infective." The "calm" strategy and "impulsive" strategy are adopted in evacuation. The emotional intensity of individual evacuation is spread by others and fluctuates, and the evacuation strategy of each individual evacuation is not single. Moreover, due to the influence of dynamic changes, the emotion of evacuation individuals affects their mobile behavior [34]. Considering that only when emotion accumulates to the time step of cell movement can pedestrian movement be driven, the improved model adopts parallel update mechanism with a unified time step Δt. In each time step Δt, the emotion is updated firstly to judge the change of the emotional state. Then, the behavioral decision-making is updated. The simulation process of this model is as follows:

① Initialize the number of people and the layout of evacuation space;

② Calculate the SFF of each cell according to equation (2);



(a) (b)

Figure 11: Target cell of "impulsive."

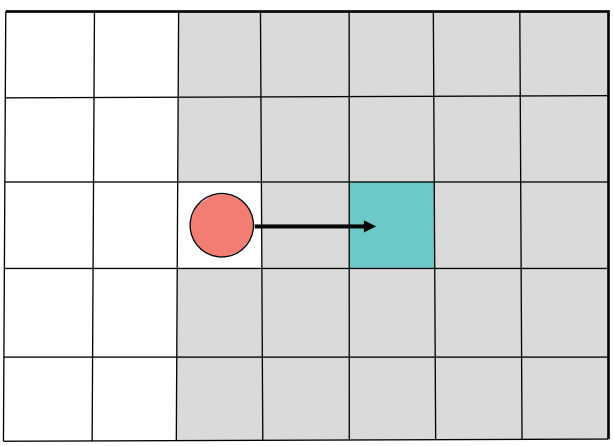


Figure 12: Density range observed for "impulsive."

③ Distribute the crowd in the evacuation space randomly and initialize the emotional intensity $PE_j(0)$ of evacuation individuals. When the emotional intensity of individual is lower than the threshold $\lambda_{emotion}$, record the location and emotional intensity of individual in the initial set S. Otherwise, record them in the initial set I;

④ Update individual emotion: calculate the emotional increment according to equations (3)–(8) and update the emotional intensity;

⑤ Update individual status: if the emotional intensity of "emotional susceptible" S is higher than the threshold $\lambda_{emotion}$, it will turn into "emotional infective" I with probability β; if the emotional intensity of "emotional infective" I is lower than the threshold $\lambda_{emotion}$, it will turn into "emotional susceptible" S with probability γ. Realize the dynamic change of evacuation crowd state, as shown in Figure 13;

⑥ Calculate the individual transfer strategy according to equations (9)–(14) and select the target cell;

⑦ Solve conflicts: when multiple evacuees choose the same target unit, there will be conflicts. In this case, an evacuation individual is randomly selected to move. For simplicity, it is assumed that the probability of each individual being selected is equal [31];

⑧ When all the evacuees in the space leave, the simulation ends; otherwise, jump to ④;

⑨ Every time the evacuation individual moves, pedometer + 1, until the end of the simulation.

4.2. Simulation Verification Analysis

4.2.1. Experiment 1. This experiment analyzes the influence of perception domain on the total evacuation time. In the improved model, the emotion perception domain of the evacuation individual is a circle with a radius R. Set a single exit room as the simulation scene. The size is 12m × 12m. The exit width is $el_1 = 3$cl. The size of each cell is cl × cl, cl = 0.4m. The initial number of people is $N = 200$, which is randomly distributed in the space and given the initial emotion value. In order to analyze the influence of perception domain radius R, set $nn = R/\text{cl}$ (cl = 0.4m). Then, $nn = 1, 2, 3, 4, 5$. Emotional intensity threshold $\lambda_{emotion}$, proportion of emotional infected people I, emotional infection coefficient β, and emotional calm coefficient γ and other parameters all affect the overall evacuation time of the system. Therefore, three different scenarios with different parameters are set for simulation experiments. Scenario 1: $\beta = 0.33$, $\gamma = 0.67$; scenario 2: $\beta = 0.5$, $\gamma = 0.5$; scenario 3: $\beta = 0.67$, $\gamma = 0.33$. Next, we analyze the impact of different perception domain radius nn on the total evacuation time under three scenarios, when the emotional threshold $\lambda_{emotion} = 0.2, 0.4, 0.6, 0.8$. Each experiment runs 100 times, and the average of the total evacuation time is taken. The simulation results are shown in Figures 14–16. It can be seen

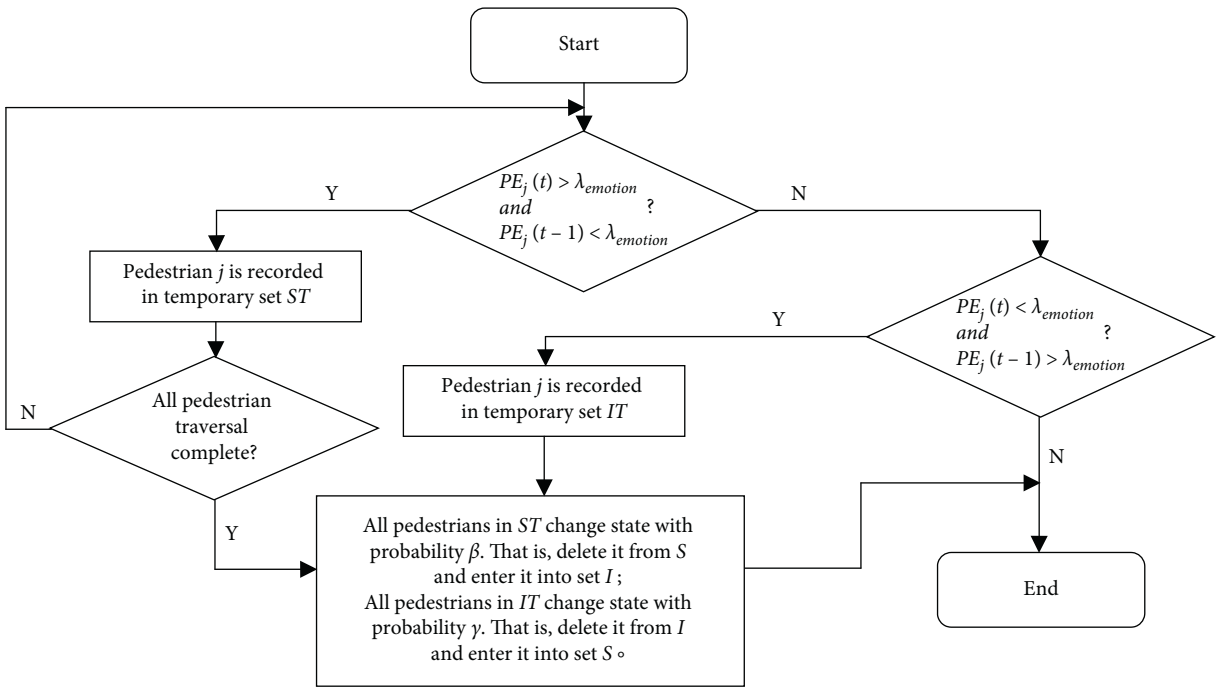


Figure 13: Update process of evacuation individual state in time t.

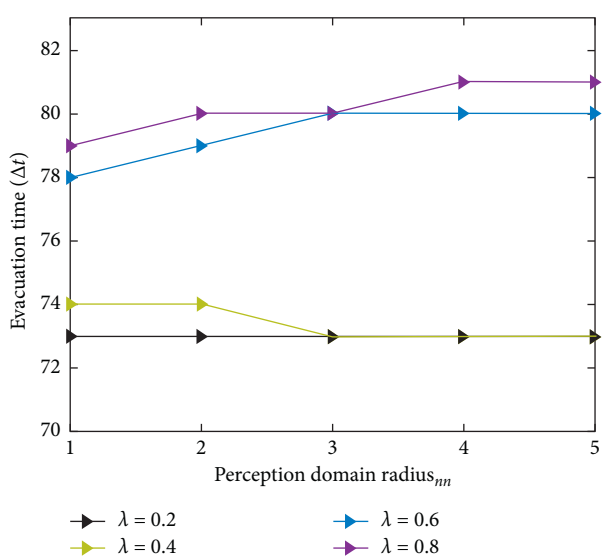


Figure 14: Total evacuation time with different perception domain radius in scenario 1 ($\beta = 0.33$, $\gamma = 0.67$).

from the figure that, when $\lambda_{\text{emotion}} = 0.2$, there is no difference in the total evacuation time between scenario 1 and scenario 3, while the total evacuation time of the perceived radius $nn = 4.5$ is smaller at scenario 2; When $\lambda_{\text{emotion}} = 0.4$, the total evacuation time of the perceived radius $nn = 4.5$ is less than $nn = 1, 2, 3$ in the three scenarios. When $\lambda_{\text{emotion}} = 0.6$ and $\lambda_{\text{emotion}} = 0.8$, the three scenarios show that the minimum total evacuation time appears when $nn = 1$, which generally shows that the larger the perception domain radius is, the larger the total evacuation time is. Therefore, when $\lambda_{\text{emotion}} < 0.5$, the larger the perceived radius nn is, the faster the total evacuation time is; when $\lambda_{\text{emotion}} > 0.5$, the larger the perceived radius nn is, the longer the total evacuation time is. The values of β and γ have no decisive influence on this conclusion. In order to verify this conclusion, we carried out simulation experiments on the same three scenarios with the initial numbers of people $N = 300$ and $N = 400$, and the results are consistent.

4.2.2. Experiment 2. The perceived domain radius of individual evacuation has a direct impact on the emotion of individual evacuation. We carried out simulation experiments on the mean system emotions in scenario 1, scenario 2, and scenario 3 under different emotion thresholds. The experimental results are shown in Figure 17. In the figure, the horizontal axis is the perception domain radius $nn = 1, 2, 3, 4, 5$. The vertical axis is the mean emotional value of the whole system, and the formula is as follows: $\text{APE}(t) = \sum_{j=1}^{N(t)} \text{PE}_j / N(t)$, that is, the cumulative sum of the mean system emotion at each time divided by the total evacuation time. The value of this parameter reflects the emotional level of the individual in the process of evacuation. The initial position and initial emotional value of the experiment were the same as that of experiment 1. The results show that, when $\lambda_{\text{emotion}} = 0.2$ and $\lambda_{\text{emotion}} = 0.4$, with the increase of perception domain radius, the mean system emotion increases; when $\lambda_{\text{emotion}} = 0.6$ and $\lambda_{\text{emotion}} = 0.8$, with the increase of perception domain radius, the mean system emotion decreases. The conclusion of this experiment is that $\lambda_{\text{emotion}} = 0.5$ can be regarded as a cut-off point of emotional value. The initial people numbers of $N = 300$ and $N = 400$ are simulated, and the conclusion is consistent.

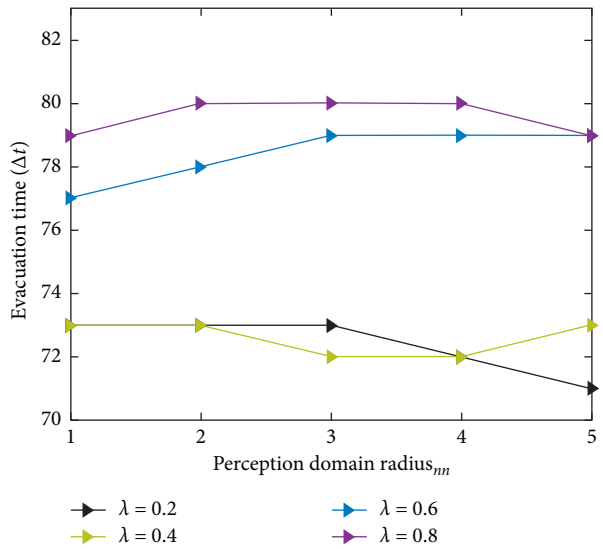


Figure 15: Total evacuation time with different perception domain radius in scenario 2 ($\beta = 0.5$, $\gamma = 0.5$).

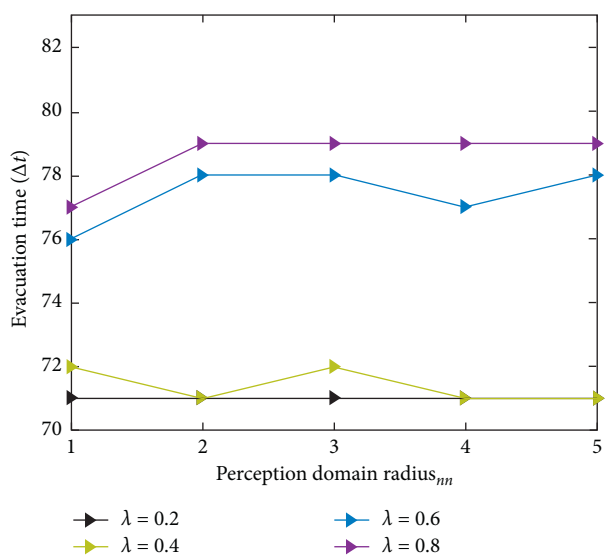


Figure 16: Total evacuation time with different perception domain radius in scenario 3 ($\beta = 0.67$, $\gamma = 0.33$).

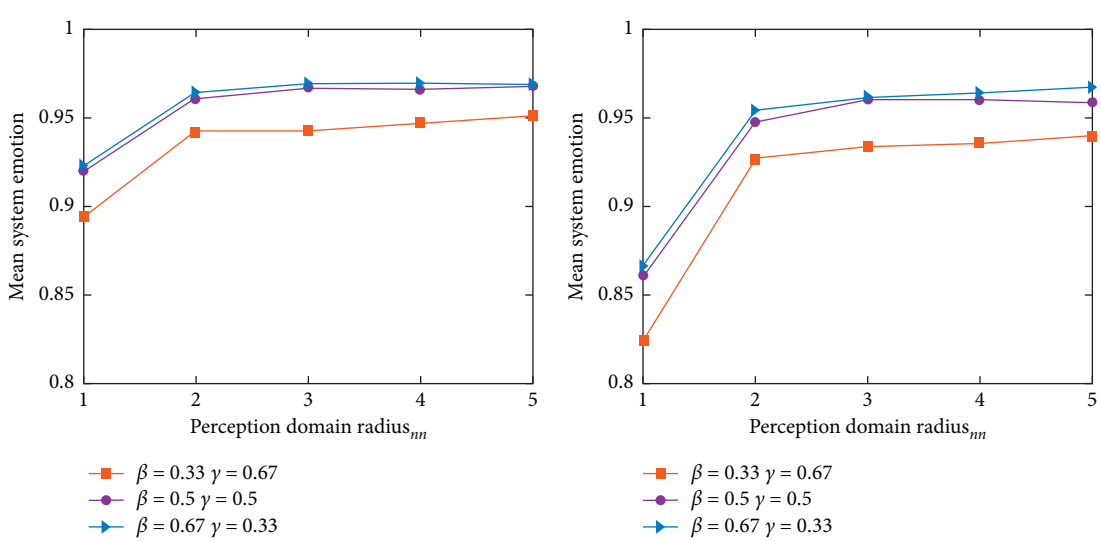


Figure 17: Continued.

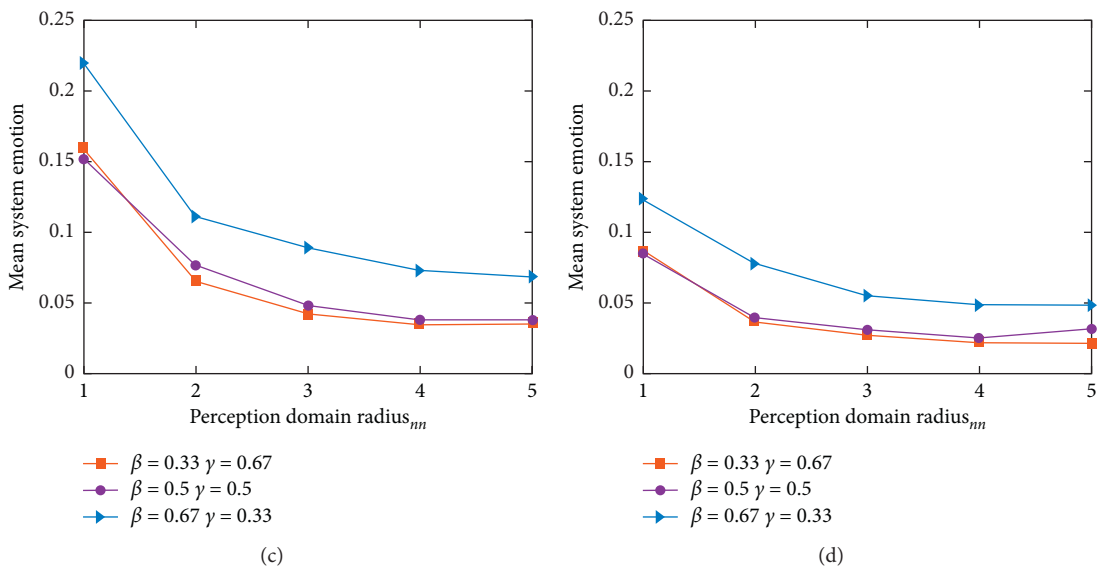


Figure 17: Mean system emotion value under different initial conditions and different perception domain radius. (a) $\lambda_{\text{emotion}} = 0.2$. (b) $\lambda_{\text{emotion}} = 0.4$. (c) $\lambda_{\text{emotion}} = 0.6$. (d) $\lambda_{\text{emotion}} = 0.8$.

4.2.3. Experiment 3. The change of emotional value will lead to the change of evacuation individual state. This experiment analyzes the influence of different perceived domain radius of evacuation individuals on the state of evacuation individuals. According to the previous two experiments, $\lambda_{\text{emotion}} = 0.5$ is the inflection point of the change of evacuation time and mean system emotional value. In order to observe the detailed changes of pedestrian status more clearly, $\lambda_{\text{emotion}} = 0.4$, $\lambda_{\text{emotion}} = 0.5$, and $\lambda_{\text{emotion}} = 0.6$ are taken in this experiment respectively, to analyze the change of the proportion of state I in the system at each time. The experiment adopts the same scene as the previous experiment. The initial position and initial emotional value are the same. The initial number of people are $N = 200$ and $N = 300$ respectively. Take $\beta = 0.5$ and $\gamma = 0.5$. The experimental results are shown in Figures 18 and 19.

It can be seen from the figure that no matter $N = 200$ or $N = 300$, the system will be transformed into group I eventually with the evacuation time when $\lambda_{\text{emotion}} = 0.4$. When the perception domain radius $nn = 1$, the change speed is slower, and when $nn = 2, 3, 4, 5$, the change speed is faster. The proportion curve of state I of perception domain radius $nn = 1$ at $\lambda_{\text{emotion}} = 0.5$ is different from other situations significantly. When $N = 200$, the proportion of I increases first and then decreases rapidly. In the last stage of evacuation, group I leaves first, and the last one is individual S. So, at the last moment, $S = 1$. When $N = 300$, the number of group I decreased slowly. This shows that for $nn = 1$, it is more beneficial to transform into group I when the number of people is small. But state group I and state group S always coexist in the system under the critical threshold no matter what kind of initial number. Under two initial numbers, when the perception domain radius $nn = 2, 3, 4, 5$, the evacuation individual status in the system is ultimately group I. From the curve change, when $nn = 2$, the change speed of I state group is slightly slower than that of $nn = 3, 4, 5$. When $\lambda_{\text{emotion}} = 0.6$, whether the initial number is $N = 200$ or $N = 300$, with the change of evacuation time, the system will be all group S. It can be seen that the state change speed of perception radius $nn = 1$ is the slowest, followed by $nn = 2$. When $N = 200$, the change speed of group I of $nn = 3$ is slightly slower than that of $nn = 4$ and $nn = 5$, while when $N = 300$; the difference of curves of $nn = 3, 4, 5$ is small. Therefore, when $\lambda_{\text{emotion}} \leq 0.5$, the larger the perceived domain radius of the evacuation individual is, the more favorable it is for the evacuation individual to change into state I. When $\lambda_{\text{emotion}} > 0.5$, the larger the perceived domain radius of the evacuation individual, the easier the evacuation individual is to change to state S.

4.2.4. Experiment 4. The improved model and CA model are compared in the same scenario to verify the effectiveness of the model. The basic scene of reference [24] is used as the simulation scene. The unit cell is cl × cl, cl = 0.4m. The single exit room size is 18cl × 14cl, and the exit width is $el_1 = 2\text{cl}$. When the initial state $t = 0$, 150 individuals were randomly distributed in the space, each occupying a cell. Moore's neighbor rule is used when evacuation individuals move. Other parameters of this model are set as $\lambda_{\text{emotion}} = 0.6$, $\xi = 0$, $\beta = 1$, and $\gamma = 1$. CA model does not consider emotional contagion, only according to the SFF rule evacuation model. After the evacuation starts, the evacuation individuals move to the exit according to different model moving rules. Each time step is 0.4s. Figures 20 and 21 record the density maps of the two models at different time steps, respectively. The cell density function is shown in equation (15). It follows Moore's neighbor rule, so num = 8 in the formula. $\text{cp}_{ji}(t)$ denotes whether there are people in the neighbor cell of cell j at time t. If $\text{cp}_{ji}(t) = 1$, the neighbor cell has people; otherwise, the neighbor cell has no people.

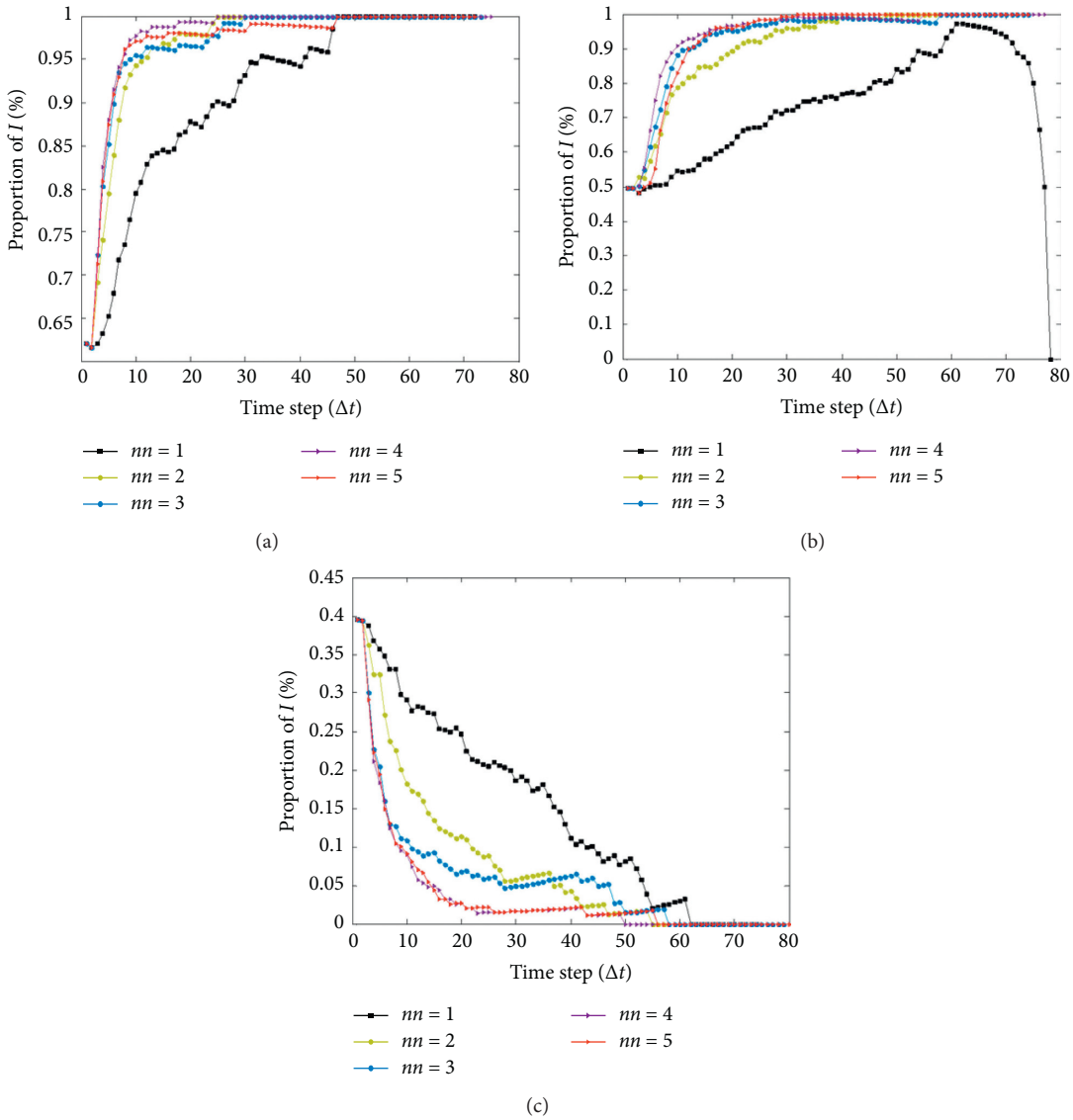


Figure 18: Proportion of state group I under different perception radius when $N = 200$. (a) $\lambda_{\text{emotion}} = 0.4$. (b) $\lambda_{\text{emotion}} = 0.5$. (c) $\lambda_{\text{emotion}} = 0.6$.

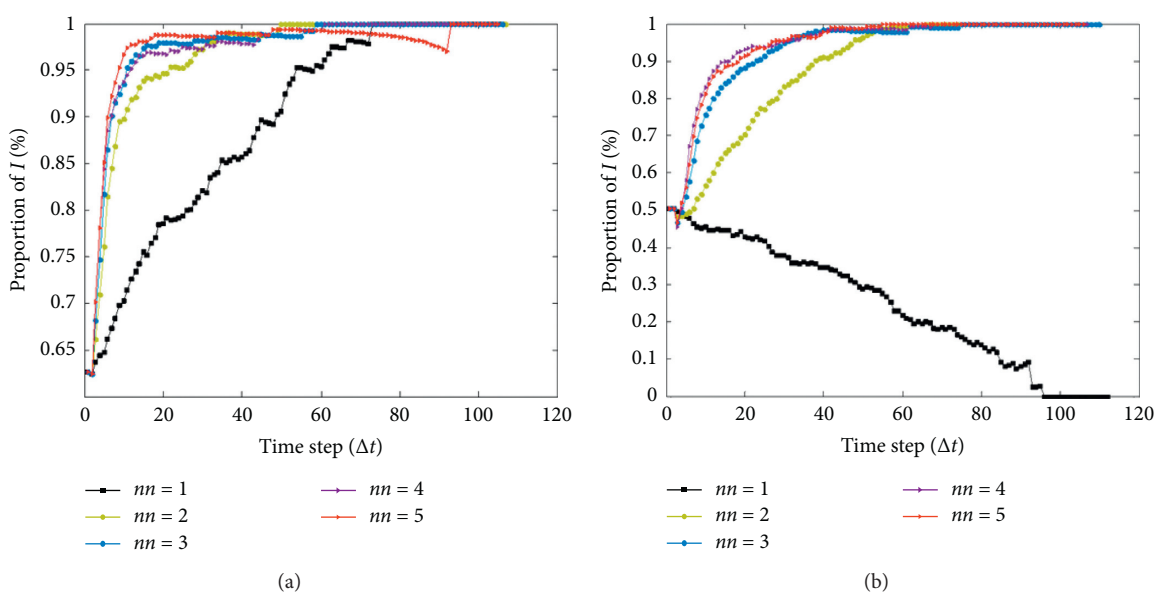
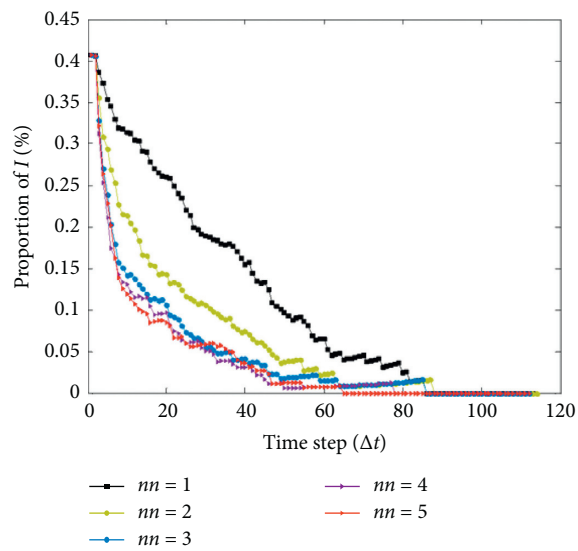
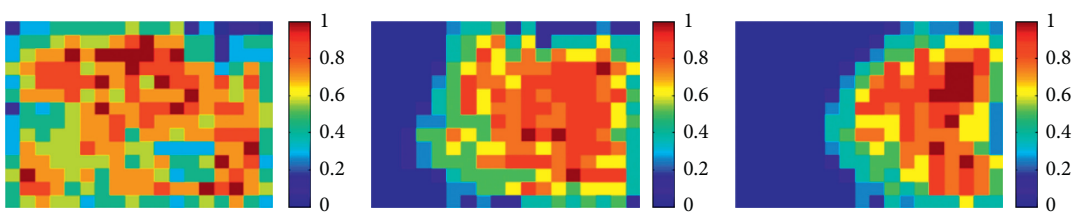


Figure 19: Continued.



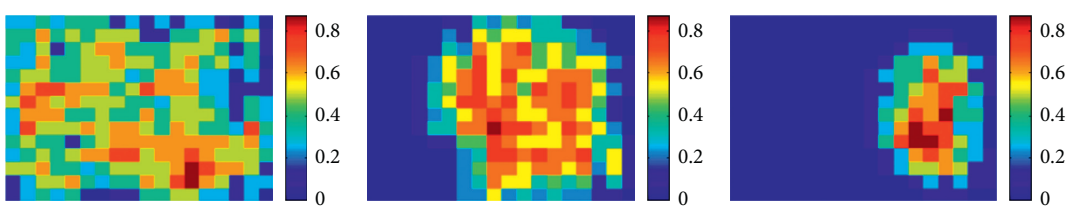
(c)

Figure 19: Proportion of state group I under different perception radius when $N = 300$. (a) $\lambda_{\text{emotion}} = 0.4$. (b) $\lambda_{\text{emotion}} = 0.5$. (c) $\lambda_{\text{emotion}} = 0.6$.



(a) (b) (c)

Figure 20: Diagram of personnel density of CA model at different times (unit: Δt). (a) $T = 0$. (b) $T = 40$. (c) $T = 80$.



(a) (b) (c)

Figure 21: Diagram of personnel density of the improved model at different times (unit: Δt). (a) $T = 0$. (b) $T = 40$. (c) $T = 80$.

$$\rho_j = \frac{\sum_{i=1}^{\text{num}} \text{cp}_{ji}(t)}{\text{num}}. \tag{15}$$

As can be seen from the figure, the personnel movement trend of the improved model is consistent with the CA model. The pedestrian flow at the exit is arched, which is in line with the characteristics of individual movement flow. In the same time step, the evacuation individuals move faster in the improved model. This is because the emotional contagion of evacuation individuals leads to the dynamic movement speed, which reduces the overall evacuation time.

4.2.5. Experiment 5. In order to further verify that the improved model can reflect the movement characteristics of individual evacuation, the scene of reference [24] is used to carry out the experiment. N people to be evacuated are distributed in a single exit room with the size of 12m × 12m randomly. Each cell is cl × cl, cl = 0.4m. That is, there are 900 cells in a single exit room. Suppose the total number of people is $N = 100, 200, 300, 400, 500, 600, 700, 800$, and the exit width is 2, 3, 4, 5, 6, 7, and 8 respectively. Comparing the total evacuation time of the improved model with that of the CA model, the simulation results are shown in Figures 22(a) and 22(b).

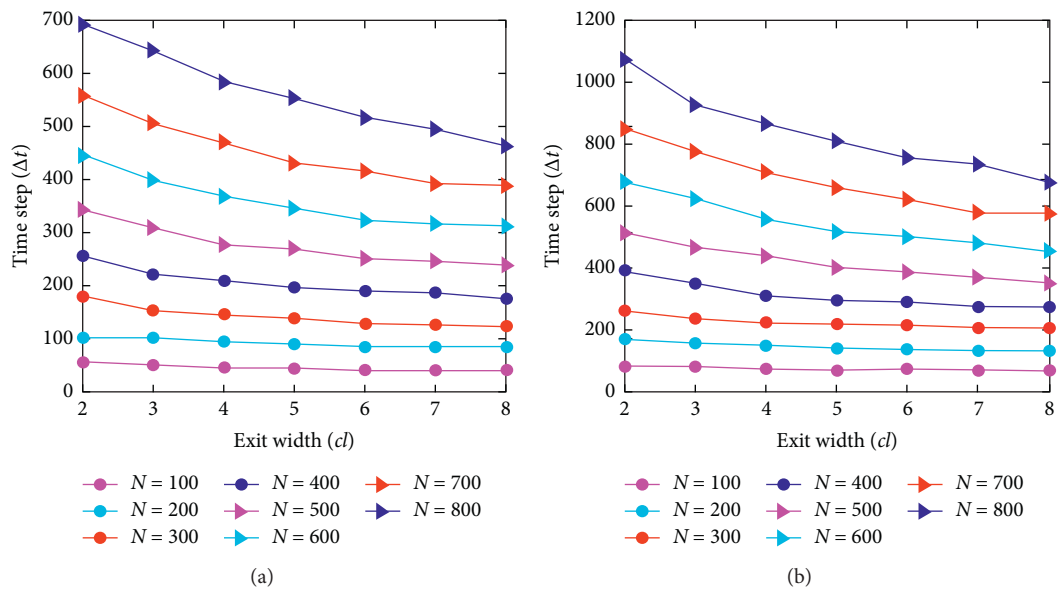


Figure 22: Total evacuation time steps of the improved model and CA model with different number of people. (a) Improved model. (b) CA model.

It can be seen from the figure that the trend presented by the two models is consistent. The wider the exit width is, the smaller the evacuation time is. The evacuation time of the improved model is shorter than that of the CA model under the same number of people and the same exit width. The reason is that the improved model takes into account the emotional contagion of individual evacuation, and different emotional groups adopt different evacuation strategies.

5. Conclusion

In this paper, the SIS model of infectious diseases is combined with cellular automata, and a crowd evacuation model considering emotional contagion is proposed. In the improved model, the dynamic movement rules of evacuation individuals are constructed. The action strategy, state transition rules, emotion update function, and linkage relationship among them are refined.

Three simulation experiments were carried out to study the influence of perception domain radius on the total evacuation time, the mean system emotion, and the proportion of state group I under different emotional threshold, infection coefficient, and calm coefficient. The results show that when the system emotion threshold $\lambda_{\text{emotion}} \leq 0.5$, the larger the perceived domain radius, the higher the mean system emotion value, the easier the evacuation individual is to change to state I and leave the site quickly. When $\lambda_{\text{emotion}} > 0.5$, the larger the perceived domain radius, the lower the mean system emotion value, and the easier the evacuation individual is to change to state S, resulting in a longer evacuation time. The conclusion can provide the following evacuation guidance for evacuation individuals. When $\lambda_{\text{emotion}} \leq 0.5$, the field of vision of the evacuation individual can be enlarged and the observation area is far away. This is conducive to judging the critical situation, forming a convergent emotion, and then make positive action strategies. When $\lambda_{\text{emotion}} > 0.5$, the evacuation individual pays attention to the emotion or behavior strategy of others nearby, which is more conducive to change personal emotion and state. If the observation scope is too large, it may lead to some misjudgments and adopt a calm strategy, resulting in prolonged evacuation time.

Two simulation experiments are compared with the CA model. The results show that the improved model can effectively simulate the process of individual evacuation, showing the characteristics of crowd conformity and exit arch. In addition, the comparative analysis also shows that emotional contagion has an impact on evacuation time. Because different emotional individuals choose different evacuation strategies, evacuation individuals move faster than the traditional CA model.

In summary, this model can not only realize the dynamic characteristics of individual state and emotion in the macroevacuation model but also analyze the microcharacteristics of individual state change and emotion fluctuation. This conclusion provides a basis for managers to grasp the characteristics of emotional contagion among people and put forward effective control strategies. Of course, the content of this chapter needs to be improved, and the simulation results need to be supported by empirical data. In the future, empirical experiments will be conducted to verify the significance of the conclusions in this chapter.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported in part by the Natural Science Foundation of Liaoning Province under Grant 2019-ZD-0558.

References

[1] H. Kim and S. Han, "Crowd evacuation simulation using active route choice model based on human characteristics," *Simulation Modelling Practice and Theory*, vol. 87, pp. 369–378, 2018.

[2] X. Z. Zheng, D. Tian, M. Zhang, C. Hu, and L. Tong, "A stairs evacuation model considering the pedestrian merging flows," *Discrete Dynamics in Nature and Society*, vol. 2019, Article ID 7615479, 2019.

[3] Y. Zhou, T. Wu, G. Zhang, and Z. Fan, "A multistory building evacuation model based on multiple-factor analysis," *Advances in Civil Engineering*, vol. 2019, Article ID 6585102, 2019.

[4] B. Liu, Z. Liu, D. Sun, and C. Bi, "An evacuation route model of crowd based on emotion and geodesic," *Mathematical Problems in Engineering*, vol. 2018, Article ID 6585102, 2018.

[5] Y. Shin, S. Kim, and I. Moon, "Simultaneous evacuation and entrance planning in complex building based on dynamic network flows," *Applied Mathematical Modelling*, vol. 73, pp. 545–562, 2019.

[6] W. Li, J. Zhu, H. Li, Q. Wu, and L. Zhang, "A game theory based on Monte Carlo analysis for optimizing evacuation routing in complex scenes," *Mathematical Problems in Engineering*, vol. 2015, Article ID 292093, 2015.

[7] Y. Peng, S.-W. Li, and Z.-Z. Hu, "A self-learning dynamic path planning method for evacuation in large public buildings based on neural networks," *Neurocomputing*, vol. 365, pp. 71–85, 2019.

[8] Y. Jiang, B. Chen, X. Li, and Z. Ding, "Dynamic navigation field in the social force model for pedestrian evacuation," *Applied Mathematical Modelling*, vol. 80, pp. 815–826, 2020.

[9] X. Guo, J. Chen, Y. Zheng, and J. Wei, "A heterogeneous lattice gas model for simulating pedestrian evacuation," *Physica A: Statistical Mechanics and Its Applications*, vol. 391, no. 3, pp. 582–592, 2012.

[10] X. Li, F. Guo, H. Kuang, Z. Geng, and Y. Fan, "An extended cost potential field cellular automaton model for pedestrian evacuation considering the restriction of visual field," *Physica A: Statistical Mechanics and Its Applications*, vol. 515, pp. 47–56, 2019.

[11] J. Wei, W. Fan, Z. Li, Y. Guo, Y. Fang, and J. Wang, "Simulating crowd evacuation in a social force model with iterative extended state observer," *Journal of Advanced Transportation*, vol. 2020, Article ID 4604187, 2020.

[12] N. Cao, L. Zhao, M. Chen, and R. Luo, "Fuzzy social force model for pedestrian evacuation under view-limited condition," *Mathematical Problems in Engineering*, vol. 2020, Article ID 2879802, 2020.

[13] M. Mitsopoulou, N. I. Dourvas, G. C. Sirakoulis, and K. Nishinari, "Spatial games and memory effects on crowd evacuation behavior with Cellular Automata," *Journal of Computational Science*, vol. 32, pp. 87–98, 2019.

[14] V. J. Blue, M. J. Embrechts, and J. L. Adler, "Cellular automata modeling of pedestrian movements," in *Proceeding of the Systems, Man, and Cybernetics*, vol. 3, IEEE, Orlando, FL, USA, October 1997.

[15] S. Liu, L. Yang, T. Fang, and J. Li, "Evacuation from a classroom considering the occupant density around exits," *Physica A: Statistical Mechanics and Its Applications*, vol. 388, no. 9, pp. 1921–1928, 2009.

[16] X. Li, Z. Geng, H. Kuang, X. Bai, and Y. Fan, "Effect of dangerous source on evacuation dynamics in pedestrian counter flow," *Physica A*, vol. 533, Article ID 122047, 2019.

[17] L. Zhang, J. Wang, Q. Shi, and T. Zang, "A new dynamic cellular automaton model for pedestrian evacuation process in stadium," in *Proceeding of the 32rd Chinese Control Conference*, IEEE, Xi'an, China, July 2013.

[18] X. Yang, B. Wang, and Z. Qin, "Floor field model based on cellular automata for simulating indoor pedestrian evacuation," *Mathematical Problems in Engineering*, vol. 2015, Article ID 820306, 2015.

[19] Y. Zhang, X. Li, N. Zhu, B. Jia, and R. Jiang, "Evacuation dynamics with smoking diffusion in three dimension based on an extended Floor-Field model," *Physica A*, vol. 507, pp. 414–426, 2018.

[20] Z. Geng, X. Li, H. Kuang, X. Bai, and Y. Fan, "Effect of uncertain information on pedestrian dynamics under adverse sight conditions," *Physica A: Statistical Mechanics and Its Applications*, vol. 521, pp. 681–691, 2019.

[21] M. Franovetter, "Threshold models of collective behavior," *American Journal of Sociology*, vol. 83, no. 6, pp. 1420–1443, 1978.

[22] R. R. McCrae, P. T. Costa, and J. S. Wiggins, *Toward a New Generation of Personality Theories: Theoretical Contexts for the Five-Factor Model," the Five-Factor Model of Personality: Theoretical Perspectives*, pp. 51–87, The Guilford Press, New York, NY, USA, 1996.

[23] D. Helbing, I. Farkas, and T. Vicsek, "Simulating dynamical features of escape panic," *Nature*, vol. 407, no. 6803, pp. 487–490, 2000.

[24] J.-H. Wang, W.-Y. Yan, Y.-R. Zhi, and J.-C. Jiang, "Investigation of the panic psychology and behaviors of evacuation crowds in subway emergencies," *Procedia Engineering*, vol. 135, pp. 128–137, 2016.

[25] M. Shi, E. W. M. Lee, and Y. Ma, "A dynamic impatience-determined cellular automata model for evacuation dynamics," *Simulation Modelling Practice and Theory*, vol. 94, pp. 367–378, 2019.

[26] X. Wang, L. Zhang, Y. Lin, Y. Zhao, and X. Hu, "Computational models and optimal control strategies for emotion contagion in the human population in emergencies," *Knowledge-Based Systems*, vol. 109, pp. 35–47, 2016.

[27] W. O. Kermack and A. G. McKendrick, "A contributions to the mathematical theory of epidemics," *Proceeding of the Royal Society of London*, vol. 115, no. 772, pp. 700–721, 1927.

[28] P. S. Dodds and D. J. Watts, "A generalized model of social and biological contagion," *Journal of Theoretical Biology*, vol. 232, no. 4, pp. 587–604, 2005.

[29] L. Fu, W. Song, W. Lv, and S. Lo, "Simulation of emotional contagion using modified SIR model: a cellular automaton approach," *Physica A: Statistical Mechanics and Its Applications*, vol. 405, pp. 380–391, 2014.

[30] C. Burstedde, K. Klauck, A. Schadschneider, and J. Zittartz, "Simulation of pedestrian dynamics using a two-dimensional cellular automaton," *Physica A*, vol. 295, no. 3-4, pp. 507–525, 2001.

[31] L. Zhang, X. Peng, L. Wang, and D. Sun, "Simulation of pedestrian evacuation considering emergency spread and pedestrian panic," *Physica A*, vol. 522, pp. 167–181, 2019.

[32] S. Li and H. Niu, "Simulation of pedestrian flow evacuation based on direction fuzzy visual field," *Journal of Transportation Systems Engineering and Information Technology*, vol. 15, no. 2, pp. 88–95, 2015.

[33] D. Nilsson and A. Johansson, "Social influence during the initial phase of a fire evacuation-Analysis of evacuation experiments in a cinema theatre," *Fire Safety Journal*, vol. 44, no. 1, pp. 71–79, 2009.

[34] T. Xu, D. Shi, J. Chen, T. Li, P. Lin, and J. Ma, "Dynamics of emotional contagion in dense pedestrian crowds," *Physics Letters A*, vol. 384, no. 3, Article ID 126080, 2019.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 6611181, 10 pages
https://doi.org/10.1155/2021/6611181

Research Article

Propagation of Uncertain Events in Multilevel Handlings at Container Terminals from the Perspective of Hypernetwork

Bowei Xu,[1] **Lingling Wang**,[1] **and Junjun Li**[2]

[1]*Institute of Logistics Science & Engineering, Shanghai Maritime University, Shanghai 201306, China*
[2]*Merchant Marine College, Shanghai Maritime University, Shanghai 201306, China*

Correspondence should be addressed to Bowei Xu; bwxu@shmtu.edu.cn

Received 31 December 2020; Revised 24 February 2021; Accepted 28 February 2021; Published 10 March 2021

Academic Editor: Jorge E. Macias-Diaz

Due to the complexity and dynamics of container terminal multilevel handlings, the occurrence of uncertain events often leads to the disruption of many operation plans and forms a chain reaction, resulting in economic losses and impacts. In this paper, the uncertain event propagation network model of container terminal multilevel handlings is constructed by using the hypernetwork theory, and the propagation influence of uncertain events between multilevel handlings at container terminals is considered. Through the topological characteristics of the uncertain event propagation network, the risk analysis is carried out for the relevance of nodes and the influence extent of uncertain events, and a simulation analysis is carried out to verify the validity of this model. The simulation results show that the model conforms to power law distribution, and the number of old nodes has an impact on the growth of hyperedge. The research provides a powerful tool for analysing the influence of uncertain events and their derivative events in the multilevel handlings at container terminals and has positive significance for reducing the losses caused by the changes of terminal plans and improving the operation efficiency.

1. Introduction

The multilevel handlings at automated container terminals are interactive each other (as shown in Figure 1). There is also "bullwhip effect" among all links in the container terminal handlings network. In fact, it is common for actual operation to deviate from the initial schedule due to various uncertain events.

The multilevel handlings at container terminals are dynamic, continuous, and diverse, but the actual handlings are complicated by various uncertain events. Uncertain events, such as ship arrival time delay, weather change, equipment failure, increase or decrease in handling capacity, and so on, will have an impact on the container terminals multilevel handlings and finally lead to plan changes. In addition, they also have chain effect. Congestion delay caused by uncertain events will be transferred among different nodes, causing other nodes to be delayed and suspended. The negative effects brought by uncertain events, such as increased time cost, idle equipment, and resources, increased container retention, delayed delivery, chaotic handling, and complex supply chain management, will also become more and more serious [1].

Focusing on the confusion of the container terminal operation plan caused by uncertain events, most scholars at home and abroad have conducted relevant research from the perspective of multilevel handlings scheduling. Tasoglu et al. [2] considered the dynamic arrival of ships and the integrated scheduling between berths and quay cranes (QCs) under random loading and unloading time and proposed a solution method based on simulation optimization. Xiang et al. [3] studied the influence of four uncertain conditions, including deviation of arrival time of ships, deviation of loading and unloading operation time of ships, unplanned arriving of ships, and failure of QCs, on the simultaneous allocation strategy of berths and QCs under discrete berths. Zhong et al. [4] considered the influence of integrated scheduling integrating QCs, automated guided vehicles (AGVs), and yard cranes (YCs). Ma et al. [5] studied the problems of unbalanced YCs works, shortage of YCs

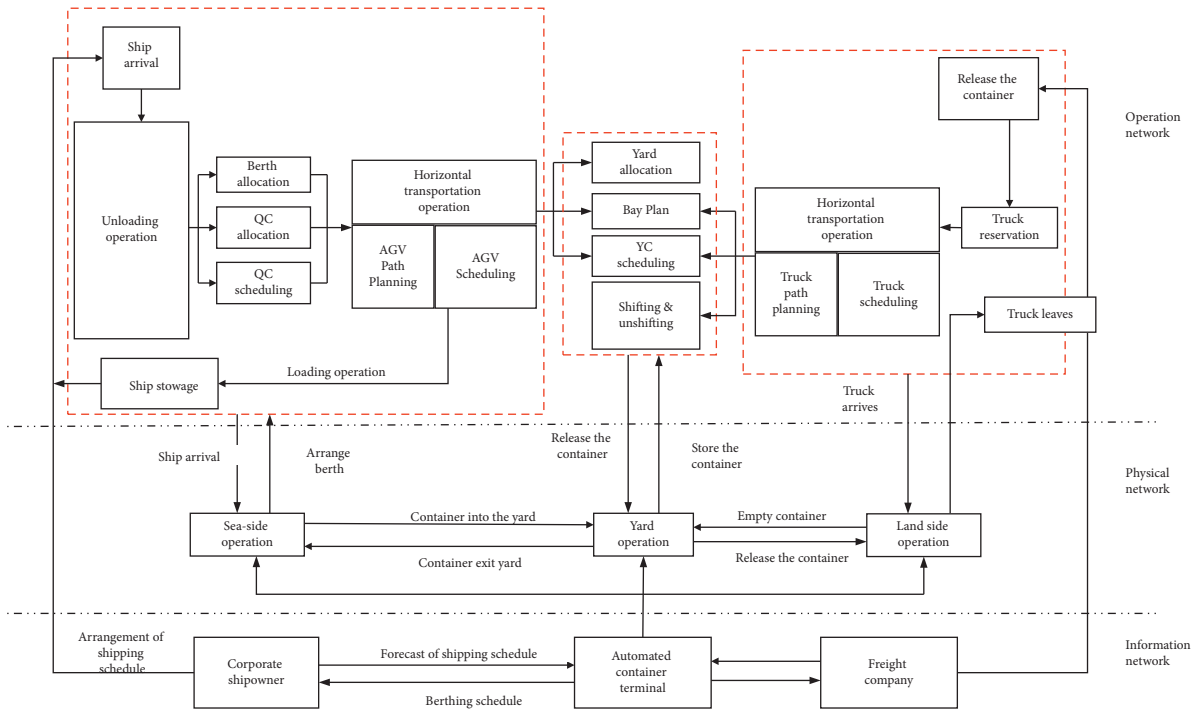


Figure 1: Schematic diagram of container terminal handlings.

resources during peak hours, and congestion of container terminals caused by random arrival time of trucks.

However, these studies did not consider the container terminal as a whole. Most studies focused on the scheduling of two or three handlings and did not consider the uncertain event propagation among the multilevel handlings at the container terminals [6–11]. Therefore, this paper takes container terminals as the study object and attempts to establish an uncertain event propagation network model based on the hypernetwork theory. The importance, relevance, and derivation of uncertain events and operation nodes are evaluated by means of the topological characteristics of hypernetwork. It helps port authorities effectively identify the risks of uncertain events, reduce the according losses, and improve the emergency response effect. The multilevel handlings at the container terminals are similar. They include the delivery/receipt of containers, the loading and unloading of yard cranes, the temporary storage of containers at the yard, the horizontal transportation of containers between berth and yard, the loading and unloading of gantry cranes, and the shipment of container vessel. Therefore, our model can be applied to different terminals. It provides a powerful tool for the correlation analysis among nodes in the propagation of uncertain events, as well as the derivation and risk assessment. The model is helpful to deeply analyse the propagation mechanism of uncertain events from the perspective of the interaction among the components of uncertain events. However, the current model only considers adding one new node per unit time in the process of hyperedge growth, and there may be multiple nodes in practice, so considering more nodes possibilities will be an improvement direction of this model.

2. Analysis of Propagation Path of Uncertain Events

2.1. Chain Propagation of Uncertain Events. The propagation effects of uncertain events and their derivative events in multilevel handlings at container terminals are mainly divided into the chain propagation and the network propagation. The chain propagation mainly describes the propagation among multilevel handlings; that is, the influence caused by an uncertain event will lead to the loss of other nodes in the multilevel handling through forward and reverse propagation. The forward propagation refers to the influence directly caused by this uncertain event, that is, the propagation direction from top to bottom along the flow sequence. As shown in Figure 2, the most direct impact caused by the bad weather is the ship delay and the suspension of handling. The ship delay will further lead to delivery delay, thus causing the supplier to bear economic losses. Reverse propagation mainly refers to the influence when an uncertain event produces a derivative event. In Figure 2, the delay of ship leads to the delay of loading and unloading operation, and the actual departure time of ship will also be delayed, and the delay will be brought to the next port. When the number of delayed ships in a container terminal reaches a certain number, it will lead to more serious port congestion, which will influence the operations of ships and container terminal. From Figure 2, the influence of uncertain events and their derivative events on multilevel handlings is mainly analysed from the longitudinal direction.

2.2. Network Propagation of Uncertain Events. The multilevel handlings at container terminals are closely connected. In addition, it often suffers diverse emergencies under complex

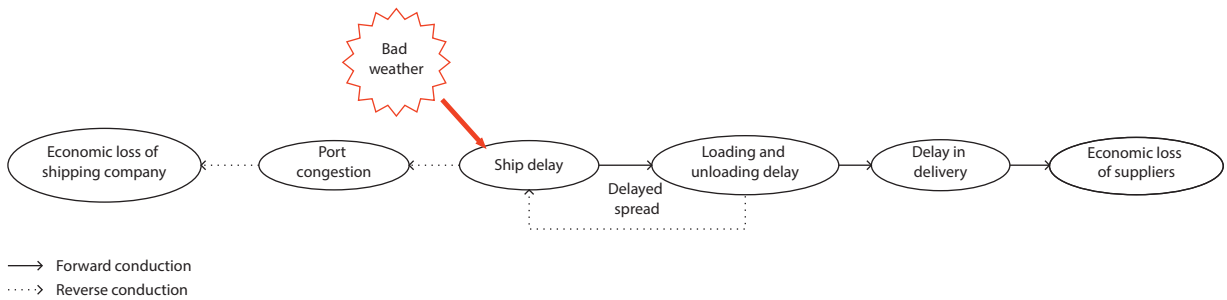


Figure 2: Example of chain propagation under sudden severe weather.

operation environment. As shown in Figure 3, when a normal node becomes a failed node after being attacked by an uncertain event, the impact of this propagation will become more and more serious over time and cause a wider spread. The collapse of a certain node will not only directly affect the superior and subordinate handlings but also affect other nodes that are not directly related but have correlations. The propagation speed and range of different uncertainties may be different. Multidimensional diffusion and superposition show strong network characteristics. These influences cross each other and overlap constantly, which makes the influence of propagation expand constantly.

3. Uncertain Event Propagation Network Model Based on Hypernetwork

3.1. Concept of Hypernetwork. The concept of hypernetwork was first put forward in 1985, and American scientist Nagurney Anna put forward that a network which is "higher but superior to the existing network" was called hypernetwork [12]. Hypernetwork is divided into three categories: hypergraph-based hypernetwork, multilayer optimized hypernetwork based on vibrational inequalities, and hypernetwork based on multisubnet integration. Hypernetwork based on hypergraph has a simple structure; it will be more suitable for describing the multistructure of the real world. Hypergraph is mainly used as a tool to connect multiple homogeneous or heterogeneous nodes through hyperedges, and thus the complex relationships among various nodes can be described concisely and accurately. This paper mainly studies hypernetwork based on hypergraphs.

The application and research of hypernetwork are mostly concentrated in the fields of knowledge learning, transportation network, balanced network, and so on [13–19]. The application of container terminals has not been tried yet. There are different levels of internal relations between different network propagation. Studying the propagation mechanism of uncertain events among container terminal handlings networks is of great positive significance to the arrangement, planning, and scheduling of container terminal multilevel handlings. It is an innovative choice to introduce hypernetwork theory into the uncertain event propagation at container terminals.

As shown in Figure 4, hypergraph-based hypernetwork can be expressed as graph $G = (V, E)$, where $V = \{V_1, V_2, \ldots, V_N\}$ is node set, Vi represents a node of G, N is the number of nodes in hypernetwork, and E represents the set of hyperedges in hypernetwork $(V_i, V_j \in V, i \neq j)$ [20].

3.2. Topological Characteristics of Uncertain Event Propagation Network Based on Hypernetwork. In this paper, based on the topological characteristics of hypernetwork, the propagation influence of uncertain events at container terminals will be described. Taking the terminal facilities as nodes and the uncertain events and their derivative events as hyperedges, each node can be connected with the hyperedge. The main topological characteristics and meanings are analysed below [21].

3.2.1. Node Degree. The degree of a node in the hypernetwork is defined as the total number of nodes connected to other nodes through hyperedge. It represents the association times between a node and other nodes and reflects the association among the operations in each link. This relevance makes different handlings interact and influence each other and provides conditions for the diffusion and propagation of uncertain events. The more association between handlings, the wider the spread of events.

3.2.2. Node Hyperdegree. The hyperdegree of a node is defined as the number of hyperedges including the nodes. In an uncertain event propagation network, it indicates the number of uncertain events affecting this node. It reflects the superposition effect of uncertain events on nodes, and the node with higher hyperdegree means that this node is easily affected by multiple uncertain events, so port authorities should improve operation flexibility and pay more attention to this kind of node in daily operation planning.

3.2.3. Hyperedge Degree. The degree of a hyperedge refers to the number of hyperedges that share a common node with this hyperedge. In an uncertain event propagation network, some uncertain events will further lead to changes of other links' handling arrangement. When the affected extent of this node exceeds a certain threshold, a new uncertain event will be formed and will lead to derivative events.

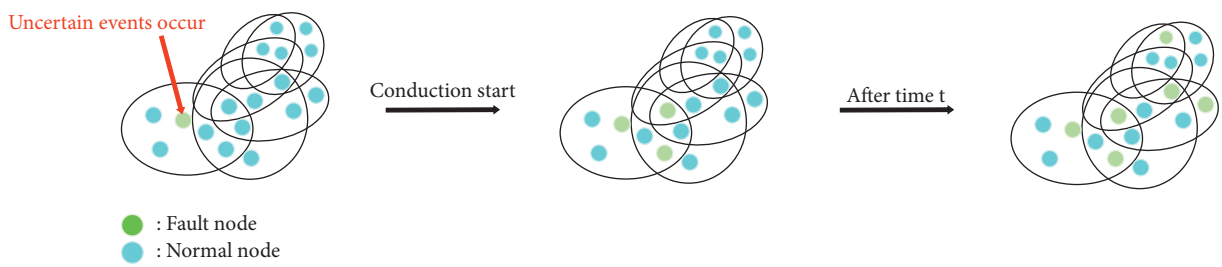


Figure 3: Example of network propagation of uncertain event.

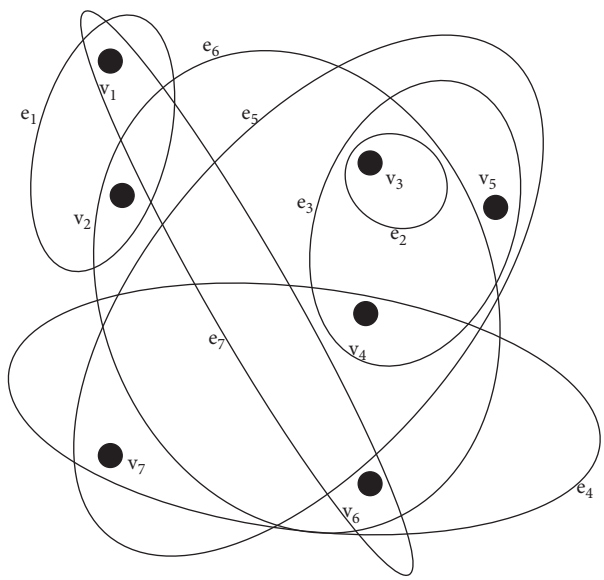


Figure 4: Graphical representation of hypergraph.

3.2.4. Hyperedge Hyperdegree. Hyperedge hyperdegree refers to the number of nodes contained in a hyperedge. In the uncertain event propagation network, it indicates the nodes' number that is affected by an uncertain event. The more nodes affected, the larger the influence range of this uncertain event. This index has important reference value for the stability and scheduling of the initial plan.

3.3. Modelling. The uncertain event propagation network model is used to describe the influence and diffusion of uncertain events in multilevel handlings at container terminal. Taking facilities as nodes and the uncertain events and their derivative events as hyperedges, a node can be affected by multiple uncertain events, and one uncertain event can also affect multiple nodes, and even some nodes will generate new derivative events after being affected. Therefore, by abutting these nodes, the propagation network model is constructed, and its schematic diagram is shown in Figure 5; the solid dots represent the old nodes, the hollow dots represent the newly added nodes, and the dashed curves represent the newly connected hyperedge.

The hypernetwork modelling algorithm is as follows.

3.3.1. Initialization. It is assumed that the hypernetwork initially has m nodes $(V1, V2, V3, \ldots, Vm)$ and a hyperedge $E1 = \{V_1, V_2, V_3, \ldots, V_m\}$ containing m nodes.

3.3.2. Hyperedge Growth. According to the process of container terminal handlings, the nodes involved in each uncertain event and its derivative events are neither a certain fixed value nor a certain equal probability, so the number of facilities (i.e., the number of nodes) is set to conform to Poisson distribution. Therefore, at time t, each time a new node V is added, a random positive integer L is produced by Poisson distribution probability:

$$p(k, \lambda) = \frac{\lambda^k}{k!} e^{-\lambda} \, (k = 0, 1, 2, \ldots), \tag{1}$$

the new node V and the $L(L \leq m)$ existing nodes in the network form a new hyperedge $Ei = \{vi1, vi2, \ldots, viL, v\}$.

3.3.3. Priority Connection. L nodes are preferentially selected from the M existing nodes in the hypernetwork according to the probability and form a hyperedge with the newly added nodes. The probability $W(i)$ of selecting the connected node i is as follows:

$$W(i) = \frac{dH(i)}{\sum_{i=1}^{n} dH(i)}. \tag{2}$$

The molecular is equal to the number of hyperedges including node i, i.e., the degree of node i. The denominator is the sum of the degrees of all current nodes.

According to the above model construction process, the propagation hypernetwork model is analysed based on the mean field theory. In the initial hypernetwork, if there are m nodes and a hyperedge, the initial degree of each node is 1. Since one node is added in the network at a time, there are $m + t$ nodes and $t + 1$ hyperedges in the hypernetwork after t times. When a new node enters the network, the old nodes are selected to form a new hyperedge, which causes the nodes' hyperdegree change. The hyperdegree of the node i should satisfy the dynamic equation:

$$\frac{\partial dH}{\partial t} = p(k; \lambda) W(i). \tag{3}$$

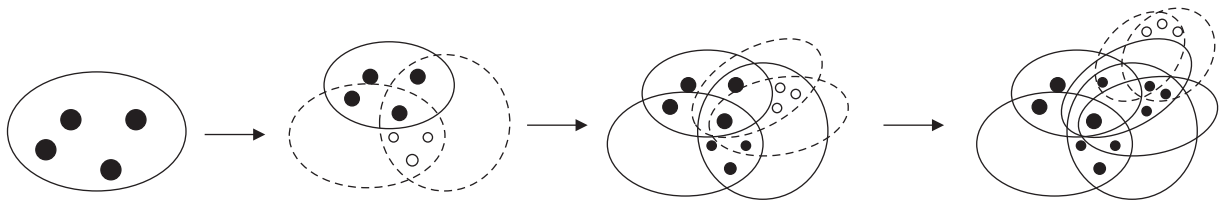


Figure 5: The evolution process of a hypernetwork.

After t times, the sum of all nodes degrees is approximately

$$\sum_{i=1}^{n} dH(i) \approx m + t(\lambda + 1). \tag{4}$$

According to Poisson distribution theorem, the expected value and variance of Poisson distribution are both λ. That is, when t is large enough, $p(k;\lambda)$ approaches to λ. Therefore, by combining formula (2) and formula (4), formula (3) can be obtained:

$$\frac{\partial dH}{\partial t} = p(k;\lambda)W(i) = \lambda \frac{dH(i)}{m + t(\lambda + 1)} \approx \frac{\lambda}{\lambda + 1} \frac{dH(i)}{t}. \tag{5}$$

When each node enters the network, the initial hyperdegree is 1, and the solution of this partial differential equation is as follows:

$$dH = \left(\frac{t}{ti}\right)^{\lambda/\lambda+1}. \tag{6}$$

Because the nodes with hyperedges in the hypernetwork are randomly selected, the probability that the nodes have hyperdegree is as follows:

$$P(dH(t) < dH) = P\left(ti > \frac{t}{dH^{\lambda+1/\lambda}}\right). \tag{7}$$

Assuming that at the same time interval, new hyperedge adding obeys uniform distribution, i.e., ti has a constant probability density $p(ti) = 1/t$, and thus substituting it into formula (7) [12], we obtain

$$P\left(ti > \frac{t}{dH^{\lambda+1/\lambda}}\right) = 1 - P\left(ti \leq \frac{t}{dH^{1+(1/\lambda)}}\right) = 1 - \frac{t}{dH^{\lambda+1/\lambda} \cdot t}. \tag{8}$$

Deriving formula (8), the hypernetwork hyperdegree distribution $P(dH, t)$ is as follows:

$$P(dH, t) = \frac{\partial P(dH(t) < dH)}{\partial dH} = \frac{\lambda + 1}{\lambda} dH^{-(2+(1/\lambda))}. \tag{9}$$

Through the derivation of formula (9), we can get the power law distribution of the node hyperdegree distribution function of the random hypernetwork constructed in this paper. And in the fourth part, the change of distribution function with different parameters n and m will be given by simulation [12].

4. Simulation Analysis

4.1. Risk Analysis of Uncertain Event Propagation Network Based on Hypernetwork. The risk analysis of uncertain events can better understand the impact of uncertain events at container terminal and help port authorities to strengthen the management of the important handling links and prevent uncertain events engendering greater impact. Therefore, the derivation and development of uncertain events can be better controlled and improved the operation efficiency [21].

In order to better explain the impact of uncertain events at container terminals, this section evaluates several types of uncertain events that occur frequently at container terminals through two indicators, the affected extent of nodes and the influence extent of uncertain events.

4.1.1. The Affected Extent of Nodes. The affected extent of nodes refers to the measurement of the derivative events that may occur after a node is affected by uncertain events. In the multilevel handling propagation network, there are many factors that may affect the initial plan, among which the main ones are as follows. (a) The number of uncertain events (N), a node may be affected by multiple events, which can be reflected by the node hyperdegree. (b) Relevance between each handling (R), which is the fundamental reason for the uncertain events to propagate in multilevel handlings. The more complex the correlation among handlings, the greater the probability of potential accidents. This index is reflected by the node degrees. (c) Derived event probability coefficient (P). Due to the different extent of association among handlings, some nodes will not continue to trigger other events, while some nodes may become new uncertain events and trigger derived events. The probability coefficient of derived events has certain randomness and can be given according to statistical data. According to the above analysis, the affected extent of nodes can be defined as $S(Vi) = Ni * Ri * Pi$.

4.1.2. The Influence Extent of Uncertain Events. The influence extent of uncertain events refers to the comprehensive influence extent of all nodes affected by uncertain events, and the derivative of events should be considered at the same time, so its influence extent of uncertain events is defined as $S(e_j) = \alpha_j \sum_{i=1}^{M} S(e_i)$.

Among them, α_j represents the coefficient of derivative events generated by uncertain events, and this index can be reflected by hyperedge. M is the number of nodes that acted by an uncertain event, and it can be reflected by the degrees of edge ej.

By sorting out the literature considering the impact of uncertain events on container terminal handlings, the uncertain events and their derivative events that may occur at container terminal handlings are shown in Table 1:

The corresponding incidence matrix $L(G)$ is as follows:

$$L(G) = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}. \quad (10)$$

Table 2 shows a final calculation result about the affected extent of nodes. It can be seen that the two nodes with the higher value are $V4$ (YC) and $V3$ (AGV). It represents that they have a higher correlation with other nodes. Therefore, it is necessary to focus on monitoring of these two nodes to reduce the occurrence of secondary and derivative events when scheduling.

Table 3 shows a final calculation result about the influence extent of uncertain events. It can be seen that the three hyperedges with the higher value are $E6$ (equipment failure), $E5$ (truck arrival delay), and $E3$ (loading and unloading operation volume). It shows that they act on more nodes and further illustrate that they have a large influence range in the whole network. Therefore, the port authorities should pay attention to improving the flexibility of the operation plan of relevant nodes so as to more effectively and timely deal with the situation that the actual handling deviates from the plan when these uncertain events occurred.

The chain propagation of multilevel handlings in Figure 6 is based on the following rule: the nodes are deleted from high to low according to the affected degree of nodes shown in Table 2. From Table 2, the highest values of node $V3$ and node $V4$ indicate that these two nodes are vulnerable to uncertain events; it is necessary to avoid the connection among these nodes and other nodes as much as possible and at the same time ensure the connectivity of the graph. However, the chain propagation graph cannot further express more complex relationships with the increase in the number of nodes. Furthermore, the chain propagation will become cumbersome. Therefore, when considering the impact of uncertainties on specific nodes in multilevel handlings at container terminals, network propagation will be more applicable.

4.2. Simulation Analysis of Uncertain Event Propagation Network Based on Hypernetwork. According to the export process of containers terminal, the freight company arranges the truck to delivery. After entering the port, the truck arrives at the designated container area and unloads. Then, the container is stored into the yard by the YC. After that, AGV will send the container to the designated QC for loading operation. In this process, if the truck was delayed on the road, the loading operation of the ship will be delayed, and the scheduled sailing time of the ship will be delayed. On the other hand, it will further delay the ship arrival at the next transit port and affect the unloading operation and other operation plan at the next port. Even if the delayed container is arranged to be transported by other ships, the delay of the container will affect the loading and unloading operation in the next port.

In this whole process, it is only the delay of the truck arrival at first, but it constantly triggered the operation of other links change, and the influence continues to expand. It even further leads to the delay of the ship arrival at the next port. Taking Yangshan Port as an example, Yangshan Phase IV Automated Terminal currently has 16 QCs, 80 YCs, and 130 AGVs, the average annual container throughput is over 4 million TEUs, and equipment investment is expected to continue increasing in the future. The propagation and evolution of uncertain events among multilevel handlings are very serious at automated container terminals. Figure 7 shows the simulation results of distribution function under different parameters n and m. The abscissa $dH(i)$ represents the node hyperdegree; hyperdegree refers to the number of hyperedges including node i, which can be used to measure the importance of node in hypernetwork. In the uncertain event propagation network, node hyperdegree indicates the number of uncertain events affecting nodes. The larger the hyperdegree is, the more vulnerable the node is. The vertical coordinate $P(dH)$ represents the hyperdegree distribution, and hyperdegree distribution $P(dH)$ reflects the level distribution that nodes be affected in the network.

According to simulation results, the probability distribution of node hyperdegree becomes smaller with the increase in node hyperdegree, and this change becomes more and more obvious with the increase in n. According to the definition of node hyperdegree, this shows that in the process of hypernetwork growth, the elder nodes are selected, the slower the growth rate of its hyperdegree distribution, i.e., the slower the growth rate of uncertain events affecting these nodes. This is because for the multilevel handling hypernetwork at container terminals, if more old nodes are selected, the overlapping probability of operation will be higher and the diffusion range of the influence caused by uncertain events will be limited. When the hypernetwork is dominated by new nodes, the interaction frequency among different operations is higher and the propagation range of uncertain events in the multilevel handlings at container terminals will be larger.

Through the above deduction, it can be concluded that the node hyperdegree distribution of the random hypernetwork constructed in this paper conforms to the power law distribution. Figure 8 shows the comparison between this hyperdegree distribution function and the power law distribution function deduced above when λ takes different values (according to the above simulation results, the date,

Table 1: List of uncertainty events and job nodes.

Uncertain events and their derivative events (Ei)	Affected job nodes (Vi)
Ship arrival delay $E1$	Berth $V1$, QC $V2$
The quantity of allocation QCs $E2$	AGV $V3$
Handling capacity $E3$	AGV $V3$, YC $V4$, yard allocation $V5$
Changes in freight demand $E4$	YC $V4$, truck $V6$, gate congestion $V7$
Truck arrival delay $E5$	AGV $V3$, YC $V4$, yard allocation $V5$, gate congestion $V7$
Equipment failure $E6$	QC $V2$, AGV $V3$, YC $V4$, truck $V6$
Weather and other factors $E7$	Berth $V1$, truck $V6$

Table 2: Analysis of the affected extent of nodes.

Node	Node hyperdegree N	Node degree R	Subhead-derived event probability coefficient P	$S(Vi)$
$V1$	2	2	0.2	0.8
$V2$	2	4	0.7	5.6
$V3$	4	4	0.7	11.2
$V4$	4	5	0.8	16.0
$V5$	2	3	0.4	2.4
$V6$	3	5	0.4	6.0
$V7$	2	4	0.5	4.0

Table 3: Analysis of influence degree of uncertain events.

Hyperedge	Hyperedge degree αj	Hyperedge hyperdegree M	$S(ej)$
$E1$	2	2	12.8
$E2$	3	1	33.6
$E3$	4	3	118.4
$E4$	4	3	104.0
$E5$	4	4	134.4
$E6$	6	4	232.8
$E7$	3	2	20.4

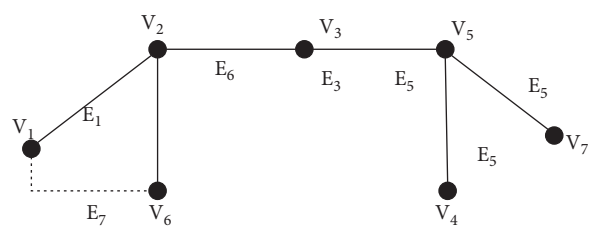


Figure 6: Chain propagation of multilevel handlings.

$n = 10000$, and $m = 5$ were chosen). According to Figure 8, when $\lambda = 4$, the two figures can basically achieve coincidence.

4.3. Comparative Analysis of Different Container Terminals. In order to further verify the rationality of the proposed model, we compare two different container terminals, Xiamen Yuanhai Terminal and Shanghai Yangshan Phase IV Terminal. Xiamen Yuanhai Terminal is equipped with 3 automatic double-trolley QCs, 16 YCs, 18 AGVs, and 8 automatic container transfer platforms, with a total of 45 multilevel handling facilities. According to the project planning, Shanghai Yangshan Phase IV Terminal is expected to be equipped with 130 AGVs, 26 QCs, and 120 YCs, with a total of 276 multilevel handling facilities. In the setting of initial network nodes, this work assumes that there are three nodes in the initial network for Xiamen Yuanhai Terminal, while Shanghai Yangshan Phase IV Terminal is set to have five nodes, and the probability distribution diagram of node hyperdegree is shown in Figure 9. The propagation of uncertain events in the multilevel handlings at container terminals depends on the connections among the interactively coupled handlings. The propagation network scale expands with the increase in the number of hyperedges and nodes.

Figure 9 shows an obvious power law attenuation trend, which indicates the uncertain event propagation networks in the multilevel handlings at container terminals have scale-free features. It also reveals two other basic characteristics of the uncertain event propagation networks: "growth" and "preferred connection," which mean that nodes with larger hyperdegree will be preferred when the influence of uncertain events is propagated. However, due to the limitation of network scale, the speed of propagation will gradually tend to be flat; that is, with the increasing hyperdegree, the distribution of hyperdegree will become smaller and smaller, as displayed in two different terminals. This characteristic of Shanghai Yangshan Phase IV network with more nodes is more obvious, as shown in Figure 9 (2), while the Xiamen Yuanhai network with less nodes will show this characteristic faster, as shown in Figure 9 (1). This is because the number of selected nodes is random when selecting nodes to form a hyperedge in each time step, which makes the number of nodes included in the hyperedge change accordingly. However, for smaller networks, due to the limitation of node's number, the number of old nodes in the hyperedge will be more and more in the later period of network growth, so it will show scale-free characteristic faster. Based on the abovementioned characteristics of hypernetwork, it also reveals that port authorities need to strengthen the maintenance of some facilities with high node

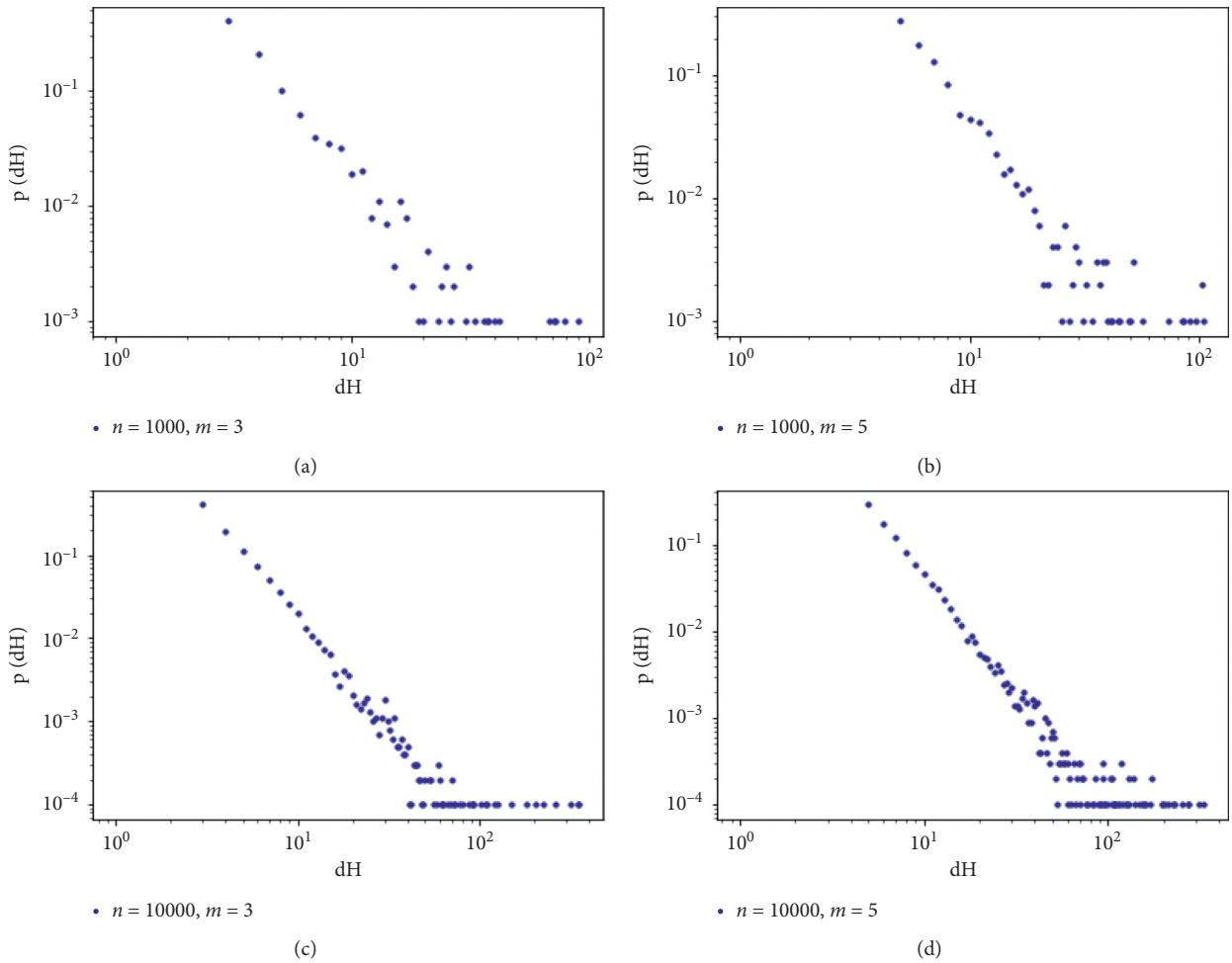


Figure 7: Double logarithmic hyperdegree distribution diagram.

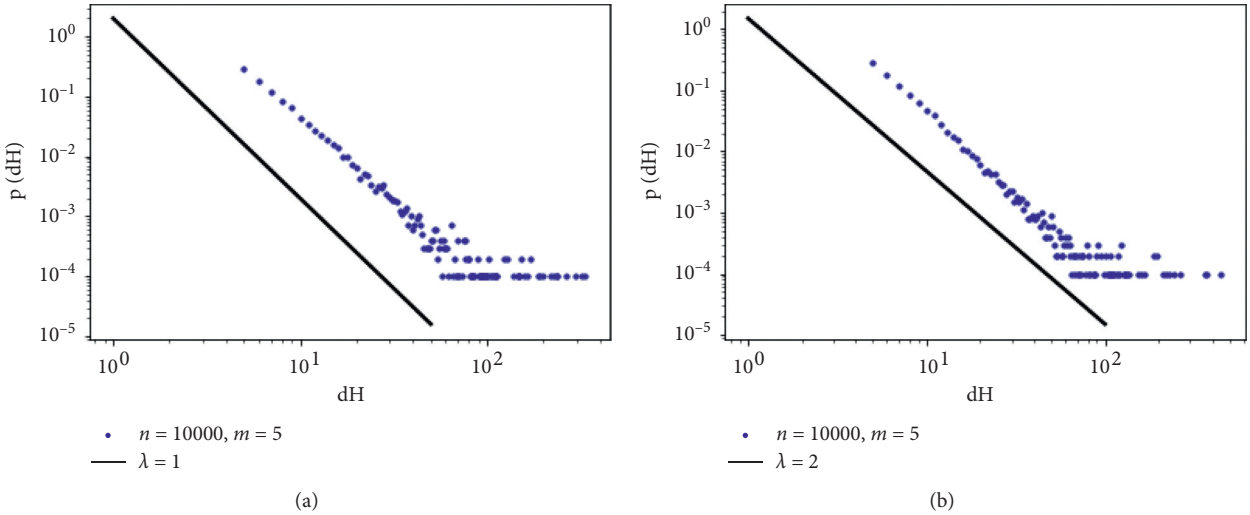


Figure 8: Continued.

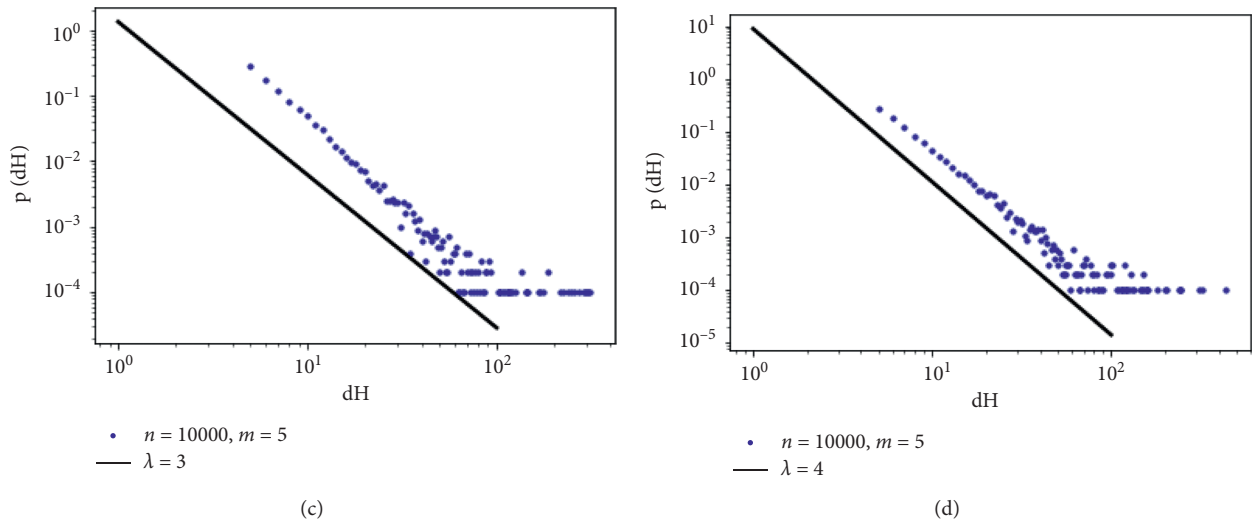


Figure 8: Simulation and theoretical analysis of double logarithmic hyperdegree distribution diagram.

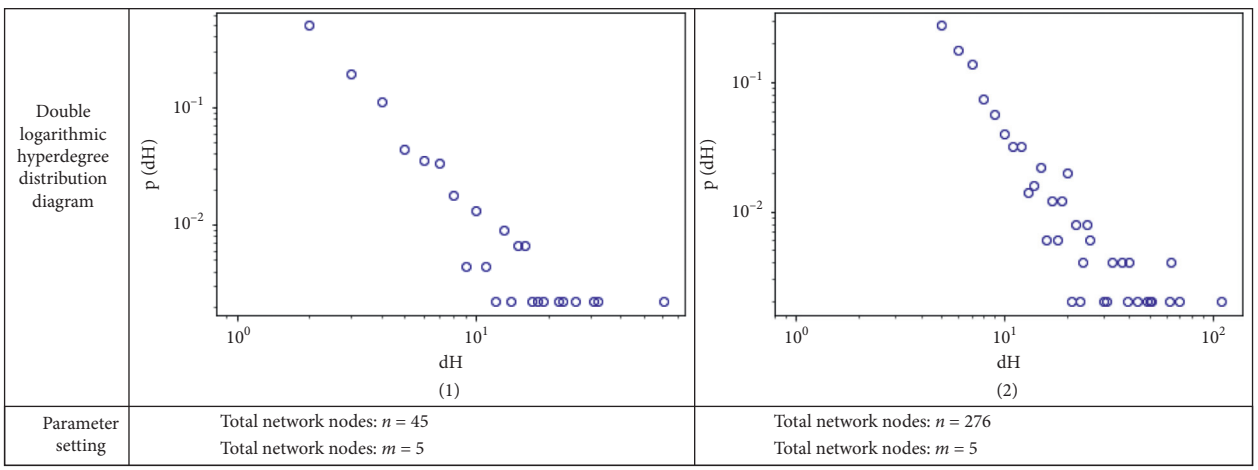


Figure 9: Hyperdegree distribution diagram of different terminals.

hyperdegree in daily operation so as to prevent such nodes from affecting more operations when uncertain events occur. At the same time, it is necessary to effectively manage equipment resources for some nodes with small hyperdegree, and this kind of resources will easily become the boundary of propagation range when uncertain events occur because nodes with high hyperdegree will use this kind of nodes to spread the influence continuously.

5. Conclusion

In this paper, the propagation influence of uncertain events in multilevel handlings at container terminals is analysed from the perspective of hypernetwork, and the container terminals are considered from the perspective of hypernetwork. At present, most of the related research studies considered uncertain events focused on a specific handling link, which could not show and analyse the correlation among operation nodes and the propagation impact of uncertain events in the whole container terminal. In view of this, this paper uses the hypernetwork theory to construct the uncertain event propagation network model in multilevel handlings at container terminals. Through the topological characteristics of the propagation network, the risk analysis is carried out, and the influence extent of some uncertain events and related nodes at container terminals are evaluated. Finally, the propagation changes of uncertain events at container terminals are analysed by simulation. The model has the following advantages:

(1) The topological characteristics of uncertain event propagation network are analysed by using the hypernetwork theory, which provides a powerful evaluation reference for analysing the correlation in the multilevel handlings and the probability of the secondary events. It is of certain value to reduce the impact of uncertain events on the initial plan arrangement, the occurrence of derivative events, and the possible losses caused by changes in the terminal plan so as to improve the handling efficiency of terminals.

(2) For the model given in this paper, by reducing or increasing the nodes and hyperedges in the network, it can form different scales' network to do a further study on the optimization of handling efficiency at container terminals.

However, the uncertain events considered in this paper are mainly a rough summary of the existing literature and other materials, and more specific situations will occur in reality, so further refining the problems and obtaining more accurate and effective analysis will be the research focus in the next stage. This paper is only a preliminary attempt to apply hypernetwork to study the propagation network of uncertain events at container terminals, and more problems such as network topology have not been discussed in depth in this paper, which will be an important direction to do a further exploring in the future.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this article.

Acknowledgments

This work was supported by the National Social Science Foundation Project of China (no. 17CGL018).

References

[1] M. R. Othman, J. Jeevan, and S. Rizal, "The Malaysian intermodal terminal system: the implication on the Malaysian maritime cluster," *International Journal of E-Navigation and Maritime Economy*, vol. 4, pp. 46–61, 2016.

[2] G. Tasoglu and G. Yildiz, "Simulated annealing based simulation optimization method for solving integrated berth allocation and quay crane scheduling problems," *Simulation Modelling Practice and Theory*, vol. 97, Article ID 101948, 2019.

[3] L. Xiang, C. Liu, and L. Miao, "Reactive strategy for discrete berth allocation and quay crane assignment problems under uncertainty," *Computers & Industrial Engineering*, vol. 126, no. DEC, pp. 196–216, 2018.

[4] O. Postolache, Y. Yang, and Y. Zhou, "Adaptive autotuning mathematical approaches for integrated optimization of automated container terminal," *Mathematical Problems in Engineering*, vol. 2019, no. 4, 14 pages, Article ID 7641670, 2019.

[5] M. Ma, H. Fan, M. Ji, and Z. Guo, "Collaborative optimization of truck delivery reservation and yc scheduling in container terminal," *Journal of Transportation*, vol. 3, pp. 202–209, 2018, in Chinese.

[6] M. Mohammadi and K. Forghani, "Solving a stochastic berth allocation problem using a hybrid sequence pair-based simulated annealing algorithm," *Engineering Optimization*, vol. 51, no. 10, pp. 1–19, 2019.

[7] T. Yu, Z. Qiang, and Z. Benfei, "A genetic algorithm based on spatiotemporal conflict between continuous berth-allocation and time-varying specific crane assignment," *Engineering Optimization*, vol. 51, no. 3, pp. 390–411, 2019.

[8] Z. Wang and C. Guo, "Minimizing the risk of seaport operations efficiency reduction affected by vessel arrival delay," *Industrial Management & Data Systems*, vol. 118, no. 7, pp. 1498–1509, 2018.

[9] H. P. Hsu, T. L. Chiang, C. N. Wang et al., "A hybrid GA with variable quay crane assignment for solving berth allocation problem and quay crane assignment problem simultaneously," *Sustainability*, vol. 11, no. 7, 2018.

[10] D. N. Prayogo, A. Hidayatno, and Komarudin, "Developing a robust optimization model for seaside operations in container terminal under uncertainty environment," *IOP Conference Series Earth and Environmental Science*, vol. 235, 2019.

[11] Y. Xie and D. P. Song, "Optimal planning for container prestaging, discharging, and loading processes at seaport rail terminals with uncertainty," *Transportation Research Part E: Logistics and Transportation Review*, vol. 119, 2018.

[12] F. Hu, *Research on the Structure, Modeling and Application of Complex Hypernetwork*, Shaanxi Normal University, Xi'an, China, 2014, in Chinese.

[13] Y. Chi, X. Tang, Y. Lian et al., "A supernetwork-based online post informative quality evaluation model," *Knowledge-Based Systems*, vol. 168, pp. 10–24, 2019.

[14] S. Kan, W. Lv, and F. Guo, "Dynamic learning super network modeling of a complex product system based on multi-organization cooperation," *Modern Physics Letters B*, vol. 32, no. 31, 2018.

[15] H. Pu, Y. Li, C. Ma et al., "Analysis of the projective synchronization of the urban public transportation super network," *Advances in Mechanical Engineering*, vol. 9, no. 6, 2017.

[16] W. Xiao and L. Wang, "The identification of maritime industry clusters using maritime industry super-network," *Ekoloji Dergisi*, vol. 107, 2019.

[17] Z. Liu, W. Chen, C. Zhang, C. Yang, and H. Chu, "Data supernetwork fault prediction model and maintenance strategy for mechanical product based on digital twin," *IEEE Access*, vol. 7, pp. 177284–177296, 2019.

[18] C. F. Zhu, Z. K. Zhang, and C. X. Ma, "Research on emergency logistics dynamic network based on super network," *Latin American Applied Research*, vol. 47, no. 1, pp. 11–16, 2017.

[19] Y. Su, J. Qin, P. Yang et al., "A supply chain-logistics supernetwork equilibrium model for urban logistics facility network optimization," *Mathematical Problems in Engineering*, vol. 2019, Article ID 5375282, 12 pages, 2019.

[20] Y. Ji, C. Wu, S. Yang, M. Guo, and Y. Zhang, "Chain evolution model of network security events based on hyper-network," *Journal of Cyber Security*, vol. 4, no. 1, pp. 89–100, 2019, in Chinese.

[21] X. Wang and L. &Gui, "Research on the evolutionary network of sudden disasters based on hypergraph," *Journal of Mathematics in Practice and Theory*, vol. 12, p. 7, 2019, in Chinese.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 5552069, 12 pages
https://doi.org/10.1155/2021/5552069

Research Article

Research on the Impact of Ecological Civilization Construction on Environmental Pollution Control in China—Based on Differential Game Theory

Liuwei Zhao[1] and Shuai Jin[2]

[1]*School of Business, Jiangsu University of Technology, Changzhou, Jiangsu 213001, China*
[2]*Computational Experiment Center for Social Science, School of Management, Jiangsu University, Zhenjiang, Jiangsu 212013, China*

Correspondence should be addressed to Shuai Jin; sjin@ujs.edu.cn

Received 11 January 2021; Revised 17 February 2021; Accepted 25 February 2021; Published 4 March 2021

Academic Editor: Stefania Tomasiello

With environmental pollution problems becoming more and more serious, administrators need to change their original system of performance achievement evaluation and build a new performance achievement view of ecological civilization. Based on the view of achievement evaluation of ecological civilization, this research develops a continuous-time differential game of environmental pollution control between local government and enterprises by taking into consideration dynamic change in pollutants in the environment into account. By analyzing the equilibrium solution of the game, the role of the assessment mechanism of ecological civilization is explored, and countermeasures are proposed to provide a scientific basis for improving the assessment mechanism of ecological civilization in China's environmental pollution control. The results have shown that enterprises' equilibrium pollutants output is negatively correlated to Ecological Civilization Index. Instead, the government's efforts to curb pollutions are positively correlated to the environmental review. With th`at, lowering enterprises' cost of production and pollution reduction alongside imposing more severe punishment would help improve pollution management. Setting up reasonable reward-punishment system would improve local governments' pollution management. Elsewhere, the research employs the method of numerical simulation to testify the relationship among the importance of environmental review, enterprises' pollutant output, and government's efforts to manage pollutions. Finally, the research also suggests necessary measures to solve environmental pollutions and introduces a numerical simulation analysis.

1. Introduction

When the environment increasingly experienced heavier environmental burden, and pollutions caused by social and economic activities have become gradually unbearable, environmental problems will become the bottleneck for the development of each country [1–3]. The contradiction between limited natural resources and the growth of human demand is one of the core issues of regional and global sustainable development [4]. The Chinese economy has prospered for over 40 years, creating a miracle of robust economic growth, but the problem of environmental pollution has worsened. The problem of environmental pollution has become an important constraint factor, hindering people's aim to improve quality of living standards. Improving environmental quality has undoubtedly played a key role in improving people's standards of living [5].

In order to improve the ecological environment, China began to launch a national-level strategy for the first time since the 10th Five-Year Plan. "Sustainable development strategy" proposed from the 15th National Congress of the Communist Party of China to the 19th National Congress of the CPC emphasized that "Lucid waters and lush mountains are invaluable assets." Environmental protection strategies are playing an increasingly important role in government management and social governance in China. Therefore, as

an important policy tool, the assessment of ecological civilization has become an important means to solve ecological environment problems and coordinate the relationship between environmental protection and economic development (ecological civilization is a new stage for civil development. That is to say, the civilization after the industrial revolution is a sum of material and spiritual civilization caused by harmony among humans, nature, and society. The harmony among nature, humans, and society, which would achieve sustainable development and virtuous cycle, should be an ideal way of living. Based on the perspective of harmony between humans and nature as well as the viewpoint of the 18th National Congress of the Communist Party of China, ecological civilization is the sum of materials, spirits, and institutions after human beings successfully protected and created ideal ecological environments. The impressive results would come from successful engineering systems of economic, political, cultural, and social construction. The results are also the measures of civilization development).

Political performance review is an important tool to help advance the modernization of the country's governance system and capacity for governance. In China's environmental governance, most scholars believe that the central government's long-term implementation of "GDP-only" performance evaluations of local officials has led to a lack of fundamental institutional incentives for environmental governance [6]. Some related issues have been discussed in existing research; for example, using the principal-agent framework, Liu et al. pointed out that the binding targets of the "eleventh five-year plan" clearly defined the administrative responsibilities of local governments and the consequences of rewards and punishments in environmental governance, and the political pressure brought by the assessment on local officials improved the intensity and effectiveness of policy implementation. However, in the process of evaluation, problems such as distorted evaluation results failure of evaluation incentives and loss of control over evaluation constraints have emerged due to asymmetric information. Most of these studies are limited to the theoretical level, and there is not much literature on the quantitative assessment of environmental performance from the perspective of game theory [7, 8].

At present, many scholars have applied the game theory method to studying the environmental pollution control problem. In real economic activities, due to the influence of the external environment, information, cognitive constraints, and other factors, rational decision-making is difficult [9]. Based on the perspective of the circular industrial chain, the stakeholders of enterprises, governments, and consumers in the resource treatment of livestock and poultry wastes are incorporated into an evolutionary game system to explore the interaction results between network cooperation subjects in the circular economy [10]. Analyzing the relationship between related stakeholders in environmental governance is essential for the model [11]. In environmental governance, enterprises are the source of environmental pollution and the main body of governance, while the government is the main body of environmental standard formulation and supervision. Therefore, there are many evolutionary game studies with the government and enterprises as the game players. By constructing the game model of government and enterprise and interenterprise and intergovernment evolution, we will study the influencing factors of local government's environmental regulation strategy [12], the influence of carbon tax and subsidy on manufacturers' low-carbon strategy choice [13], the strategic choice behavior of enterprises and environmental protection departments, and causes of enterprise pollution, prevention, and control focus [14]. Benhong et al. establish an evolutionary game model between government regulatory departments and electronic and electrical products' manufacturing enterprises and to carry out simulation analysis under different strategies [15]. Benhong et al. studied from the perspective of enterprises that the establishment of a cooperative mechanism for resource sharing in the ecosystem can effectively promote value cocreation, thereby improving the environmental performance of enterprises [16].

Now, the use of differential game method to discuss environmental pollution control is at the initial stage. Yeung established a cooperative differential game model for cross-border industrial pollution. The main characteristics of the model are as follows: various industrial enterprises are competing with each other, and the governments cooperate with each other in the treatment of environmental pollution, and for the first time, they obtain a time-consistent solution in the cooperative differential game of pollution control, and for the first time, a time-consistent solution was obtained in the cooperative differential game of pollution control [17]. Kossioris and Plexousakis pointed out that the dynamic problems of pollution and resource management need to be analyzed with the framework of differential game, and the nonlinear feedback Nash equilibrium solution of lake pollution control is obtained [18]. Finally, the feedback Nash equilibrium solutions, optimal control solutions, and open-loop Nash equilibrium solutions are compared. Yeung used the method of differential countermeasures to explore the game between policymakers and enterprises that pursued maximum profits while causing pollution. In this game, the government decided the level of pollution tax collection, and the enterprise determined the output level, obtained feedback Nash equilibrium solution, and extended the game to multiple enterprises. Akihiko Yanase established a differential game model for international pollution control [19]. The study points out that the level of environmental policy determined deviates from the social optimal level in the non-cooperative policy game, and the equilibrium result of emission tax game is worse than that of command-control game. This is because the emission tax will cause more pollution and a decline in social welfare.

Due to the special attributes of environmental indicators, the academic literature on environmental performance appraisal focuses on the following three aspects: first, the analysis and research on the practice and exploration of environmental performance appraisal; second, the establishment of environmental performance appraisal indicator system, including the principles of index construction, the selection and design of indicators, and the determination of

weights; and third, the government's environmental performance appraisal system research, including the determination of the object of assessment, the standard of assessment criteria, and the selection of assessment methods. However, most of these studies are limited to the theoretical level, and there is not much literature on the quantitative assessment of environmental performance from the perspective of game theory.

Environmental pollution control is a continuous dynamic game process. Differential game is a dynamic game model that studies the competition and cooperation of two or more parties in a continuous tablet. Therefore, based on the differential game model, this paper constructs a game model of environmental pollution control between the government and enterprises under the constraint of the evaluation mechanism of ecological civilization. The following issues are studied:

First, how does the introduction of the assessment mechanism of ecological civilization construction interfere with the dynamic evolution of the pollutant stock?

Second, how does this dynamic alter the efficient setting of the punishment level for a given level of emissions?

Third, what strategies would come out when the central and local information is asymmetric in the evaluation of ecological governance performance of local government?

This paper is organized as follows. In Section 2, we construct a continuous-time differential game of environmental pollution control between local government and enterprises by taking into consideration dynamic change in pollutants in the environment into account. Impact analysis of feedback Nash equilibrium of the system is discussed in Section 3. In Section 4, a numerical illustration serves to illustrate (i) how dynamic state variables shape expectations and variance in the stock of pollutants and (ii) how related and key parameter values affect equilibrium strategies of both sides of the game. In Section 5, the policy recommendations for this study are given. Finally, the conclusions and limitations of this study are given.

2. Methodology

2.1. Problem Characteristics and Basic Assumptions. Due to the existence of conflicts of interest in economic development and environmental protection, enterprises environmental pollution control has typical game characteristics, which also makes game theory one of the important tools for studying ecological environmental governance [20, 21]. For the convenience of research, it is assumed that, within a continuous-time $s \subset [t_0, t]$, there is a superior government supervision department, a local government supervision department, and n polluting enterprises in a certain area. n polluting enterprises will carry out a certain amount of sewage in any time s, which will have a direct impact on the dynamic system of the ecological environment. The pollution of the ecological environment will cause damage to enterprises and the government. The damage cost is a linear function of the level of pollutants in the ecological environment. It is assumed that the polluting enterprises are "economic man" with limited rationality. In order to pursue the maximum profit, it will risk the selection of environmental violations such as discharge pollutant without permission, beyond pollution limits, and other illegal activities. If $q_i(s)$ represents the output level of enterprise i within time s, the yield of $q_i(s)$ will also produce $e_i(s) = \gamma q_i(s)(\gamma > 0)$ pollutants. The polluting enterprise obtains income through production and operation, and the amount of pollutants is related to the output of its production. Therefore, the polluting enterprise's production gain and production cost can be written as a function of pollutant $e_i(s)$. Under the supervision of the government, polluting enterprises i will take measures to control the generated pollutants, and its cost function C_i is a concave function, namely, $C_i'' > 0$. Similarly, the government will also deal with environmental pollution, and its cost function C_g is also a concave function, namely, $C_g'' > 0$. The superior government will give certain rewards and punishments according to the local government's ecological civilization and environmental protection performance. The size of the rewards and punishments is a linear function of the ecological civilization and environmental performance. Finally, it is assumed that the government and the polluting enterprises have the same discount rate ρ.

2.2. Environmental Quality Dynamic Equation. The environmental quality dynamic equation [22, 23] is as follows:

$$\begin{aligned} \dot{E}(s) &= m\left(P_{\max} - E(s)\right) + \sum_{j=1}^{n}\left[e_j(s) - \mathrm{Re}_j(s)\right] - \beta\mu(s), \\ E(s_0) &= E_0 = 200, \end{aligned} \tag{1}$$

where $E(s)$ represents the amount of pollutants in the ecosystem. Dynamic changes in environmental quality $\dot{E}(s)$ mainly depends on three variables: the total amount of pollutants discharged by n polluting enterprises, the amount of pollutants that local governments have eliminated through pollution control, and the amount of pollutants that the ecosystem system absorbs through its self-circulation. $e_j(s)$ represents the amount of pollution generated by the polluting enterprises j during the production and the operation process, but not necessarily all of them are discharged into the environment; R_j $(0 < R < 1)$ represents the purification rate of pollutants generated during the production and the operation process by the enterprise j's technological innovation; $R_j e_j(s)$ represents the amount of pollutants purified by the enterprise's technological innovation; $\sum_{j=1}^{n}[e_j(s) - \mathrm{Re}_j(s))]$ indicates the total amount of pollutants discharged by the polluting enterprises n; $\beta\mu(s)$ indicates the elimination of the amount of pollutants in the eco-environment system through local government efforts; $\beta(\beta > 0)$ indicates the rate of pollutants eliminated by local governments in the fight against pollution; $mE(s)$ is the pollution flow absorbed by the ecological environment

system during the self-circulation process; and m indicates the self-purification rate of the ecological environment system. It is assumed here that the self-absorption capacity of the ecological environment system is a decreasing function of the pollution stock. That is to say, the heavier the pollution is, the worse the environmental self-purification ability could be; $P_{\max}$ indicates that there is a maximum level of pollution, and P_0 represents the initial level of pollutants in the eco-environment system at time t_0.

2.3. Enterprise Objective Function. The objective function of the enterprise is as follows:

$$\int_{t_0}^{t} \left\{ \begin{array}{l} a_i e_i(s) - \frac{v_i}{2}(e_i(s))^2 + k[e_i(s) - R_i e_i(s) - \overline{e}_i] - \frac{C_i}{2}(R_i e_i(s))^2 \\ -bE(s) - G(e_i(s) - R_i e_i(s)) - \alpha(s) g[e_i(s) - R_i e_i(s) - \overline{e}_i]\omega \end{array} \right\} e^{-\rho(s-t_0)} \mathrm{d}s. \tag{2}$$

The income of the enterprise can be divided into two parts: one is the net income brought by the production and operation of the enterprise, and the other part is the income from the excessive discharge of the enterprise. The net income of the enterprise is the production utility minus the production cost. According to the hypothetical net income, it is related to $e_i(s)$, which is $a_i e_i(s) - (v_i(e_i(s))^2/2)$, where a_i and v_i represent the utility and cost coefficients, respectively, and $a_i > 0$, $v_i > 0$. $k[e_i(s) - R_i e_i(s) - \overline{e}_i]$ indicates the additional benefits that the excess pollutants bring to the enterprise i, where $\overline{e}_i$ is the maximum amount of pollutants that the environmental protection department permits and $k(k > 0)$ indicates the coefficient of return that the unit exceeds the standard amount of pollutants.

The cost and expenses of the enterprise include the following: the damage cost to the enterprise caused by environmental pollution $bE(s)(b > 0)$, the sewage charges paid by the enterprise, the cost of abatement, and the fines paid after the company exceeds the standard discharge. The cost and expenses of the enterprise include the following: the damage cost to the enterprise caused by environmental pollution, the emission charges paid by the enterprise, the cost of emission reduction, and the fines paid after the enterprise exceeds the standard discharge. According to Article 12 of the Regulations on the Administration of the Collection and Use of Emission Charges, which was implemented by the People's Republic of China in 2003, pollutants discharged into the environment shall be paid in a quantity according to the amount and G shall be recorded as the cost of unit discharge; $G(e_i(s) - R_i e_i(s))$ indicates the emission charges paid by the enterprise i to the local government. $((R_i e_i(s))^2 C_i)/2$ $(C_i > 0)$ is the cost of pollution control required by the enterprise i to reduce the amount of $R_i e_i(s)$ pollutants. $\alpha(s) g[e_i(s) - R_i e_i(s) - \overline{e}_i]\omega$ indicates the amount of fines paid by the enterprise i after the government found out that the pollutants were discharged; $\alpha(s)g$ indicates the possibility of the enterprise's excessive discharge of pollutants, where g is the level of government's ecological civilization and environmental performance and $\alpha(s)$ is the importance of ecological civilization and environmental performance assessment (the higher the importance, the more the local government pays more attention to the environment, and the stronger the supervision) calculates the fine to be paid by the company exceeding the standard of pollutant discharge. ω represents the unit penalty to be paid by the enterprise whose exceeds the standard of pollutant discharge quantity.

2.4. Government Objective Function. The objective function of the government is as follows:

$$\int_{t_0}^{t} \left\{ \begin{array}{l} \sum_{j=1}^{n}\left[a_j e_{ij}(s) - \frac{v_j}{2}\left(e_j(s)\right)^2\right] + \sum_{j=1}^{n} \mathrm{G}\left(e_j(s) - \mathrm{Re}_j(s)\right) - \frac{C_g}{2}(\beta\mu(s))^2 \\ + \sum_{j=1}^{n} \alpha(s) g\left[e_j(s) - \mathrm{Re}_j(s) - \overline{e}_j\right]\omega + \frac{\vartheta\alpha(s)}{2}\left[g\mu(s) - g_0\mu_0(s)\right] - \zeta E(s) \end{array} \right\} e^{-\rho(s-t_0)} \mathrm{d}s. \tag{3}$$

The government's income mainly consists of four parts: the material welfare created by the enterprise production and operation for the society, the pollutant discharge paid by the enterprise, the fines paid after the enterprise exceeds the standard pollutant discharge, and the rewards given by the superior government. The government's losses and costs are composed of three parts: $\beta\mu(s)(\beta > 0)$ is the damage to the local people caused by environmental pollution, the other is the governance cost that the government spends to manage the environment $C_g(\beta\mu(s))^2/2$ $(C_g > 0)$, and the third part is the punishment given by the unqualified of the superior government.

$\sum_{j=1}^{n}[a_j e_{ij}(s) - v_j(e_j(s))^2/2]$ indicates the material welfare created by the enterprise's production and operation for the society, and it is also the GDP growth part created for the country; $\sum_{j=1}^{n} G(e_j(s) - \mathrm{Re}_j(s))$ represents the income from the enterprises' pollutant charges to the government; $\vartheta\alpha(s)[g\mu(s) - g_0\mu_0(s)]/2$ $(\vartheta > 0)$ represents the superior government rewards and punishments for local governments, where g_0 and $\mu_0(s)$ represent the standard values of ecological civilization achievements and environmental governance efforts formulated by the superior government to local governments, respectively. When the product of local government's environmental governance efforts and ecological civilization construction $g\mu(s)$ is greater than the critical value $g_0\mu_0(s)$, $\vartheta\alpha(s)[g\mu(s) - g_0\mu_0(s)]/2$ represents the reward given by the superior government to the local government; on the contrary, when the product of local government's environmental governance efforts and ecological civilization construction $g\mu(s)$ is less than the critical value $g_0\mu_0(s)$, $\vartheta\alpha(s)[g\mu(s) - g_0\mu_0(s)]/2$ represents the punishment given by the superior government to the local government.

2.5. Model Solution. In order to make equation (1) have a unique continuous solution $E(s)$, first, we need to construct a set of bounded, continuous, and differentiable value functions $V_c(E)$ and $V_g(E)$. Therefore, the Hamilton–Jacobi–Bellman (HJB) equation is constructed as follows:

$$\begin{aligned}\rho V_c(E) = \operatorname*{Max}_{e_i(s)\geq 0} &\left\{\begin{array}{c} a_i e_i(s) - \frac{v_i}{2}(e_i(s))^2 + k_i[e_i(s) - \mathrm{Re}_i(s) - \bar{e}_i] - G(e_i(s) - \mathrm{Re}_i(s)) \\ -bE(s) - \frac{C_i}{2}(\mathrm{Re}_i(s))^2 - \alpha(s)g[e_i(s) - \mathrm{Re}_i(s) - \bar{e}_i]\omega \end{array}\right\} \\ &+ V_c'(E)\left\{\begin{array}{c} m(P_{\max} - E(s)) - \beta\mu(s) \\ + \sum_{j=1}^{n}[e_j(s) - \mathrm{Re}_j(s))] \end{array}\right\},\end{aligned} \tag{4}$$

$$\begin{aligned}\rho V_g(E) = \operatorname*{Max}_{e_i(s)\geq 0} &\left\{\begin{array}{c} \sum_{j=1}^{n}\left[a_j e_{ij}(s) - \frac{v_j}{2}(e_j(s))^2\right] + \sum_{j=1}^{n}\alpha(s)g[e_j(s) - \mathrm{Re}_j(s) - \bar{e}_j]\omega - \zeta E(s) \\ + \sum_{j=1}^{n} G(e_j(s) - \mathrm{Re}_j(s)) + \frac{\vartheta\alpha(s)}{2}[g\mu(s) - g_0\mu_0(s)] - \frac{C_g}{2}(\beta\mu(s))^2 \end{array}\right\} \\ &+ V_g'(E)\left\{\begin{array}{c} m(P_{\max} - E(s)) - \beta\mu(s) \\ + \sum_{j=1}^{n}[e_j(s) - \mathrm{Re}_j(s))] \end{array}\right\}.\end{aligned} \tag{5}$$

First, equation (4) finds the first-order partial derivative with respect to $e_i(s)$ and obtains the following equation:

$$\begin{aligned} &a_i - v_i e_i(s) + k_i(1-R) - G(1-R) - C_i\mathrm{Re}_i(s) \\ &\quad - \omega\alpha(s)g(1-R) + V_c'(E)(1-R)n = 0. \end{aligned} \tag{6}$$

After simplification, we can get the following equation:

$$e_i(s) = \frac{a_i + (k_i - G - \omega\alpha(s)g + V_c'(E)n)(1-R)}{v_i + C_i R^2}. \tag{7}$$

Next, equation (5) finds the first-order partial derivative with respect to $\mu(s)$. After simplification, we can get the following equation:

$$\mu(s) = \frac{g\vartheta\alpha(s) - 2\beta V_g'(E)}{2C_g\beta^2}. \tag{8}$$

In order to obtain a linear value function, let $V_c(E) = l_1 + h_1 E(s)$ and $V_g(E) = l_2 + h_2 E(s)$, where l_1, l_2, h_1, and h_2 are all constants. Then, substituting $V_c'(E) = h_1$, $V_g'(E) = h_2$

into equations (4) and (5), the following equation is obtained:

$$\rho\left(l_1 + h_1 E(s)\right) = a_i e_i(s) - \frac{v_i}{2}\left(e_i(s)\right)^2 + k_i\left[e_i(s) - \mathrm{Re}_i(s) - \overline{e}_i\right] - bE(s) - G\left(e_i(s) - \mathrm{Re}_i(s)\right) - \frac{C_i}{2}\left(\mathrm{Re}_i(s)\right)^2 - \alpha(s) g\left[e_i(s) - \mathrm{Re}_i(s) - \overline{e}_i\right]\omega + h_1\left\{m\left(P_{\max} - E(s)\right) + \sum_{j=1}^{n}\left[e_j(s) - \mathrm{Re}_j(s)\right)\right] - \beta\mu(s)\right\}. \tag{9}$$

$$\rho\left(l_2 + h_2 E(s)\right) = \sum_{j=1}^{n}\left[a_i e_i(s) - \frac{v_i}{2}\left(e_i(s)\right)^2\right] + \sum_{j=1}^{n}\left[\mathrm{G}e_j(s) - \mathrm{Re}_j(s)\right] - \frac{C_g}{2}(\beta\mu(s))^2 + \sum_{j=1}^{n}\alpha(s) g\left[e_j(s) - \mathrm{Re}_j(s) - \overline{e}_j\right]\omega + \frac{\vartheta\mu(s)}{2}\left[g\mu(s) - g_0\mu_0(s)\right] - \zeta E(s) + h_2\left\{m\left(P_{\max} - E(s)\right) + \sum_{j=1}^{n}\left[e_j(s) - \mathrm{Re}_j(s)\right)\right] - \beta\mu(s)\right\}. \tag{10}$$

From equation (9), we can know that the following equation is established:

$$\rho h_1 E(s) = -bE(s) - h_1 m E(s). \tag{11}$$

By simplification of equation (11), the following equation can be obtained:

$$h_1 = \frac{-b}{\rho + m}. \tag{12}$$

Similarly, we can get

$$\rho h_2 E(s) = -\zeta E(s) - h_2 m E(s), \qquad h_2 = \frac{-\zeta}{\rho + m}. \tag{13}$$

As a result, we get

$$V_c'(E) = h_1 = \frac{-b}{\rho + m}, \tag{14}$$

$$V_g'(E) = h_2 = \frac{-\zeta}{\rho + m}. \tag{15}$$

Substituting equations (14) and (15) into (7) and (8), the simplification can be obtained as a feedback Nash equilibrium strategy $[e_i^*(s), \mu^*(s)]$ in the following form:

$$e_i^*(s) = \frac{a_i(\rho + m) + (1 - R)((\rho + m)(k - G - \omega\alpha(s) g) - bn)}{\left(v_i + C_i R^2\right)(\rho + m)}, \tag{16}$$

$$\mu^*(s) = \frac{g\vartheta\alpha(s)(\rho + m) + 2\zeta\beta}{2(\rho + m) C_g \beta^2}. \tag{17}$$

3. Impact Analysis of Feedback Nash Equilibrium

(1) The equilibrium pollutant production of enterprises $e_i^*(s)$ is negatively correlated with the importance of ecological civilized environmental performance assessment $\alpha(s)$, and equation (16) finds the first-order partial derivative with respect to $\alpha(s)$; we can get descent gradient $(\omega g(1 - R)/(v_i + C_i R^2))$. By partial guidance of the equilibrium pollutant production $e_i^*(s)$ on the importance of the assessment of ecological civilization and environmental performance $\alpha(s)$, we can obtain $(\partial e_i^*(s)/\partial\alpha(s)) = (-\omega g(1 - R)/(v_i + C_i R^2))$. Since $0 < R < 1$ and ω, g, v_i, and C_i are all greater than zero, $(-\omega g(1 - R)/(v_i + C_i R^2)) < 0$ is always established. Therefore, it can be explained that the equilibrium pollutant production of the enterprise $e_i^*(s)$ is negatively correlated with the importance of the ecological civilization and environmental performance assessment $\alpha(s)$. The descent gradient is $(\omega g(1 - R)/(v_i + C_i R^2))$; that is, when the superior government raises the importance of assessing the ecological civilization and environmental protection performance, and the local government will pay more attention to environmental protection performance. On the one hand, the local governments will take more effective measures to control the discharge of pollutants; on the other hand, the supervision of the pollutant enterprises will continue to strengthen, which will help curb enterprise discharge pollutant without permission, beyond pollution limits, and other illegal activities.

(2) The government's equilibrium environmental governance efforts $\mu^*(s)$ are positively related to the

importance of ecological civilization and environmental performance assessment $\alpha(s)$, and equation (17) finds the first-order partial derivative with respect to $\alpha(s)$; we can get ascent gradient $(g\vartheta/2C_g\beta^2)$. By partial guidance of the equilibrium environmental governance efforts $\mu^*(s)$ on the importance of the assessment of ecological civilization and environmental performance $\alpha(s)$, we can obtain $(\partial\mu^*(s)/\partial\alpha(s)) = (g\vartheta/2C_g\beta^2)$. Since g, ϑ, C_g, and β are all greater than zero, then $(g\vartheta/2C_g\beta^2) > 0$ is always established. This shows that the government's equilibrium environmental governance efforts $\mu^*(s)$ are monotonously increasing function, and the importance of ecological civilization and environmental performance assessment will promote the government's equilibrium environmental governance efforts, so as to obtain better environmental pollution control effects.

(3) The equilibrium pollutant production of enterprise $e_i^*(s)$ is negatively correlated with the unit of pollutant charge G, and equation (16) finds the first-order partial derivative with respect to G; we can get descent gradient $(1 - R/(v_i + C_iR^2))$. With partial guidance of the equilibrium pollutant production $e_i^*(s)$ on the pollutant charge G, we can obtain $(\partial e_i^*(s)/\partial G) = (-(1 - R)/(v_i + C_iR^2))$. Since $0 < R < 1$ and v_i and C_i are all greater than zero, then $(-(1 - R)/(v_i + C_iR^2)) < 0$ is always established. Therefore, we can know that when the government raises the amount of the unit of pollutant charge, the production and operation cost of the enterprise will increase, and the enterprise will take measures to control the pollutant discharge, which can play a role in restraining the pollution output of the enterprise.

(4) The reduction in the production and operation costs and environmental treatment costs of enterprises can all improve the pollution control effect of ecological civilization and environmental performance assessment. The reduction in emissions costs of enterprises has led to the descent gradient increase in the equilibrium pollutant production of enterprises $(\omega g(1 - R)/(v_i + C_iR^2))$. It shows that when the cost of emissions of enterprises is reduced, the pollution control effect of the government's ecological civilization and environmental performance assessment is further strengthened.

Similarly, reduction in production and operation costs v_i of enterprises has led to the descent gradient increase in the equilibrium pollutant production of enterprises $(\omega g(1 - R)/(v_i + C_iR^2))$. It shows that when the production and operation costs of enterprises are reduced, the pollution control effect of the government's ecological civilization and environmental performance assessment will be further strengthened.

(5) Increasing the penalties ω for the environmental illegal activities of polluting enterprises can improve the effect of environmental pollutant control. By partial guidance of the equilibrium pollutant production $e_i^*(s)$ on the penalties ω, we can obtain $(\partial e_i^*(s)/\partial\omega) = (-\alpha(s)g(1 - R)/(v_i + C_iR^2))$. Since $0 < R < 1$ and $\alpha(s)$, g, v_i, and C_i are all greater than zero, then $(-\alpha(s)g(1 - R)/(v_i + C_iR^2)) < 0$ is always established. So, we see that the enterprises' equilibrium pollutant production $e_i^*(s)$ is negatively related to the punishment ω from the government. That is to say, the punishment for the discharge pollutant without permission, beyond pollution limits, and other environmental illegal activities of enterprises will be continuously increased. When the number of fines for the discharge pollutant without permission, beyond pollution limits, and other environmental illegal activities of enterprises is greater than the cost of environmental pollution control, it may reduce the environmental illegal activities of enterprises.

(6) The superior government will increase the rewards and punishments ϑ for the local government's ecological civilization and environmental protection achievements, which can improve the level of its environmental governance efforts. By partial guidance of the government's equilibrium environmental governance efforts $\mu^*(s)$ on the rewards and punishments ϑ, we can obtain $(\partial\mu^*(s)/\partial\vartheta) = (g\alpha(s)/2C_g\beta^2)$. Since $\alpha(s)$, g, C_g, and β are all greater than zero, then $(g\alpha(s)/2C_g\beta^2) > 0$ is always established. So, we see that the government's equilibrium environmental governance efforts $\mu^*(s)$ is positively related to the rewards and punishments ϑ. Therefore, the superior government should increase the reward and punishment of environmental performance assessment; when the environmental performance assessment really affects the government's interests, it will have more power to invest in environmental pollution control.

(7) The government's lower environmental governance costs C_g can raise the level of its equilibrium environmental governance efforts $\mu^*(s)$. The government's environmental pollution control cost $C_g$is reduced, making the rising gradient $g\vartheta/2C_g\beta^2$ of the government's equilibrium governance effort $\mu^*(s)$ related to the importance of ecological civilization and environmental performance assessment $\alpha(s)$ increase. It shows that, with the reduction in government environmental governance costs, the government's equilibrium pollution control efforts will further increase.

4. Numerical Examples

The results presented thus far, while analytically compelling, may further benefit from practical grounding in a real-world example. This section thus assesses the results of the game theoretic models against a realistic backdrop of the

assessment mechanism of ecological civilization construction deployed at Hubei Province, China. A numerical illustration serves to illustrate (i) how dynamic state variables shape expectations and variance in the stock of pollutants and (ii) how related and key parameter values affect equilibrium strategies of both sides of the game.

4.1. Hubei Province, China. Hubei Province is in the south of China and in the middle reaches of the Yangtze River. It is between 29°05′ and 33°20′ north latitude and 108°21′ to 116°07′ east longitude and boasts advantageous geographical location. The Yangtze River runs from west to east and traverses the province's 1062 kilometers. The Yangtze River and its largest tributary, the Han River, run the river, the water network is vertical and horizontal, and the lake is densely covered. Hubei Province is also known as the "Thousand Lakes Province."

After Hubei province included ecological civilization construction in the performance review, the local government has implemented stringent environmental management systems. Since 2016, the Hubei provincial government has outlawed 1103 piers, recovered coastline by 143 km and cleared 744 solid waste storage points. Therefore, the proportion of category I~III water quality in 27 primary tributaries of Yangtze River has been rising year by year. Moreover, the local government has enhanced lake governance by removing 1.2754mn acres of fences and breeding cages, outlawing 274.5k acres land of fertilizer (manure) breeding as well as pearl farming. Hubei's government has also completed renewal of sewage management plant in 88 towns. It also spent RMB$30bn to deploy sewage management facilities in all of its towns and relocated 12784 livestock and poultry farms in the breeding forbidden zone. The funding was raised via local government bonds. From January–October 2018, water quality of Tongshun River's tributaries was improved from category V (poor quality) to category IV. The water quality of Shending River, Ju River, and Yiai River has experienced continual improvement (https://baijiahao.baidu.com/s?id=1632754276131852877).

The following model is combined with the previous model to assign values to external variables and give numerical simulations (numerical simulation is through numerical calculation and image display method to achieve the purpose of research on the control of environmental pollution behaviors of enterprises). First, assume that there are four enterprises in total, that is, $n = 4$ (the number of enterprises can be set to any positive integer and the adjustment of parameters will have a certain effect on the results of numerical simulation but does not affect the overall trend). The enterprise obtains certain benefits through production and operation activities. For the convenience of calculation, suppose the benefit coefficients is $a_i = 20$; at the same time, it costs a certain amount, and the initial cost coefficients are $v_i = 0.1$. For the enterprises, they need to pay a certain sewage charge to the government for each the unit of pollutants discharge, if the initial unit of sewage charges is $G = 0.4$. The enterprise conforms to the hypothesis of "rational economic man." In order to pursue the maximum profit, it will risk the selection of environmental violations such as discharge pollutant without permission, beyond pollution limits, and other illegal activities and assume that the income per unit of excess pollutant discharge will bring $k = 10$ to the enterprise. The government imposes a certain fine on the enterprise's discharge pollutant without permission, beyond pollution limits, and other illegal activities if the penalty for exceeding the pollutants per unit is $\omega = 2$. The government's supervision is influenced by the importance of ecological civilization and environmental performance assessment $\alpha\ (s)$; the value range of $\alpha\ (s)$ is $(0, 1)$, and the value of $\alpha\ (s)$ is closer to 1, explaining that the ecological civilization environmental performance assessment is more important and the initial $\alpha(s)$ is set to 0.8. The government's environmental performance level g is in the range of $0 < g < 1$, and the value of g is closer to 1, explaining that the government's environmental performance is higher and the initial g is set to 0.7. Therefore, under the supervision of the government, the enterprises will carry out a certain degree of technical input to improve the pollution purification rate. The value of the pollution purification rate R in the production process caused by the change in technology is $0 < R < 1$, and the initial R is set to 0.2. The cost coefficient of environmental management of enterprises is $C_i = 0.3$. When the ecological environment is polluted, the government will curb environmental pollution, and suppose the amount of pollution $\beta = 0.3$ is being eliminated by the government on curb environmental pollution. The cost coefficient of government's curb environmental pollution is $C_g = 0.04$. The superior government will give certain rewards according to the efforts of the local government's curb environmental pollution and the level of environmental performance, and the initial coefficient of reward is $\vartheta = 0.01$. The damage to the local people caused by environmental pollution $\zeta = 0.1$ and the harm to the enterprise $b = 0.05$, the discount rate of the enterprise (government) $\rho = 0.005$, and the self-purification rate of the ecological environment $m = 0.1$. The maximum level of pollution $P_{\max} = 1000$.

4.2. Analysis of Pollutant Stock and Sensitivity Analysis of Equilibrium Feedback. Substituting the above-set parameter values into the feedback Nash equilibrium solution of the differential game, we can conclude that $e_i^*\ (s) = 225.54$ and $\mu^*\ (s) = 80.14$. Substituting the value of the feedback Nash equilibrium solution into the environmental quality dynamic equation (1), the following equation can be obtained: $E(s) = 797.869 - 597.869e^{-0.1s}$. The dynamic evolution of environmental quality is shown in Figure 1. From Figure 1, we can see that as time s goes on, the environmental quality $E(s)$ will continue to improve and eventually stabilize. This can also better explain the rationality and high practical value of the model. At the beginning of time, the environmental quality has improved rapidly with the external intervention for a period of time, but the longer the time, the external the environmental governance and scientific and technological means to regulate the environmental behavior of enterprises, which makes the amount of pollution and

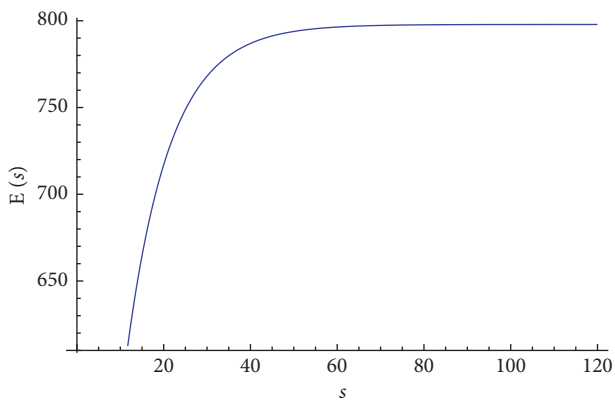


Figure 1: The dynamic evolution of environmental quality $E(s)$.

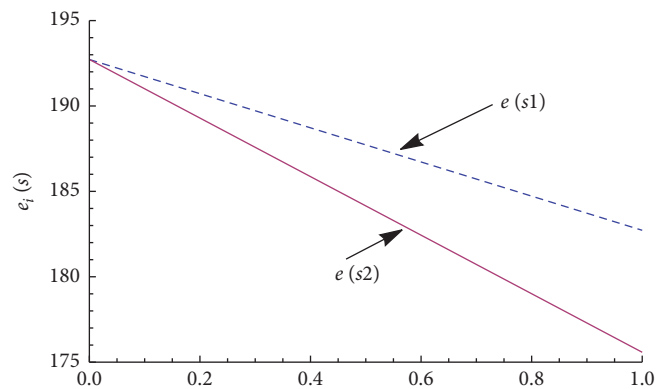


Figure 2: Evolution trend of enterprise pollutant production $e_i(s)$ under different regulations.

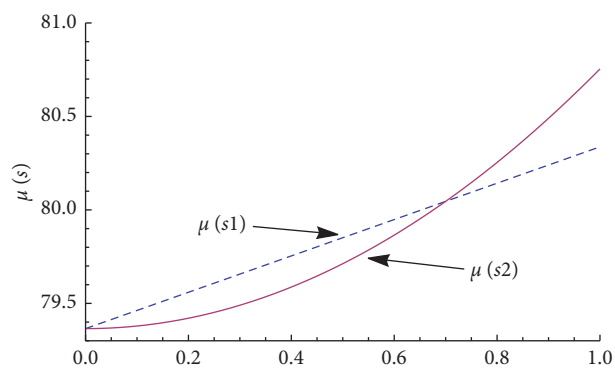


Figure 3: The evolutionary trend of government environmental governance efforts $\mu(s)$ under the influencing factors.

environmental self-absorption reach a balance so that the environmental quality will evolve to a stable state.

In Figure 2, based on the selected parameter values in Section 4.1, the environmental performance assessment $\alpha(s)$ is used to locate the independent variables, the blue imaginary curve $e(s1)$ is obtained using Mathematica 9.0 software, and it indicates the evolution trend of the pollutant output of the enterprise under the condition that the importance level of the ecological civilization environmental performance assessment $\alpha(s)$ changes; in the same way, the environmental performance assessment $\alpha(s)$ and unit sewage charges G are used to locate the independent variables, and the red solid curve $e(s2)$ is obtained using Mathematica 9.0 software; it indicates the evolution trend of pollutant production of enterprises under the dual role of unit sewage charges and the importance of ecological civilization and environmental performance assessment $\alpha(s)$. The production level of enterprise pollutants $e_i(s)$ is negatively related to the importance level of ecological civilization and environmental performance assessment. With the increasing importance of ecological civilization and environmental performance assessment, the output of enterprise pollutants has been declining. However, if the means of increasing the unit sewage charges is taken at the same time, the production of pollutants will be reduced at a faster rate, as shown by the curve $e(s2)$.

In Figure 3, based on the selected parameter values in Section 4.1, the environmental performance assessment $\alpha(s)$ is used to locate the independent variables, and the blue imaginary curve $\mu(s1)$ is obtained using Mathematica 9.0 software; it represents the evolutionary trend of the government's level of environmental governance efforts as a result of change in the importance level of ecological civilization and environmental performance assessment $\alpha(s)$; in the same way, the environmental performance assessment $\alpha(s)$ and the government's environmental performance level g are used to locate the independent variables, and the red solid curve is obtained using Mathematica 9.0 software, and it represents the evolutionary trend of the government's environmental governance efforts under the dual role of the importance of ecological civilization and environmental performance assessment $\alpha(s)$ and the government's environmental performance level g. It can be seen that the level of government's environmental governance efforts $\mu(s)$ is positively related to the importance of ecological and environmental protection performance assessment $\alpha(s)$. By comparing the curve $\mu(s1)$ and the curve $\mu(s2)$, we can find that the slope of the curve $\mu(s2)$ increases the slope of the curve larger than the slope of the curve $\mu(s1)$ with the double effect of the importance of ecological and environmental protection performance assessment and the government's environmental performance level, indicating the assessment of ecological civilization and environmental performance. The greater the effect, the greater the intensity of punishment, the better the performance of government environmental governance efforts $\mu(s)$.

5. Policy Implications

This study develops the differential game method to construct a game model of environmental pollution control between enterprises and local governments accounting for uncertain dynamics in pollution stocks and environmental factors and obtain a set of feedback Nash equilibrium solutions. By the analysis of the equilibrium solution and the numerical simulation, governmental ecological civilization and environmental protection efforts could reduce enterprises' discharged pollution. The government's efforts to curb environmental pollution could also be improved.

Therefore, we must fully recognize that ecological civilization environmental assessment plays a role in guiding, stimulating, and restraining environmental protection. By the analysis of the equilibrium solution, with validation via a numerical simulation based on the Hubei case, this study returns several insights.

5.1. Improve the Weight of Ecological Civilization Indicators in the Government Performance Appraisal $\alpha(s)$ and Formulate Reasonable Local Government Ecological Civilization Environmental Protection Performance Appraisal and Reward and Punishment System ϑ. As mentioned above, we can assume that the equilibrium pollution production of enterprises $e_i^*(s)$ is negatively correlated with the importance of ecological civilized environmental performance assessment $\alpha(s)$; the government's equilibrium of environmental governance efforts $\mu^*(s)$ is positively related to the importance of ecological civilization and environmental performance assessment $\alpha(s)$. Therefore, it is necessary to enhance the importance of ecological civilization indicators $\alpha(s)$ in the government's performance evaluation and effectively revoke the concept that GDP is the only standard for local development and government work performance.

Establish a reasonable ecological civilization and environmental performance evaluation mechanism to ensure the fairness of the assessment. On the one hand, it is necessary to conduct regular environmental performance assessments, such as monthly reports, semiannual assessments, and annual assessment mechanisms. On the other hand, environmental performance assessment can be used as a management tool. This can not only stimulate active attitudes and work behavior of relevant personnel but also may have a drumming effect on their negative work attitudes and behavior [24]. Therefore, we must fully consider the fairness of performance appraisal. We can refer to the three-dimensional model of performance appraisal fairness proposed by Levy and Williams [25]. We should design the environmental performance appraisal system from the distribution fairness dimension, program fairness dimension, and interactive fairness dimension of assessment.

Construct a reasonable ecological civilization and environmental protection performance reward and punishment mechanism. Local governments' interests have different manifestations according to different classification criteria. From the content point of view, the government interests mainly include political interests and economic interests. From the perspective of hierarchy, it includes three levels: interests of individual officials, interests of local governments and departments, and the overall interests of government organizations. The personal interests of officials lie in the fact that each official as an economic person and a social person has his/her own value orientations. Local government and departmental interests refer to them as a strong and relatively independent interest subjects, pursuing different levels of institutional interests, namely, specific sectoral interests. Finally, the overall interests of the organization include the rapid growth of government public expenditure, the expansion of institutions, and the conscious inertia of institutions [26]. Therefore, the construction of ecological civilization and environmental protection performance reward and punishment mechanism need to start from the above aspects: (a) link between the political interests and economic interests of local governments, (b) and link between personal interests of officials, local government interests, and (c) departmental interests and the overall interests of government organizations.

5.2. Accelerate the Establishment of Reward and Punishment Mechanism for Corporate Emissions and Establish Reasonable Standards for Collecting Sewage Charges. Formulate reasonable reward and punishment mechanism for the enterprise environmental behavior. Firstly, we recommend increase in penalties for the enterprises making pollution based on the environmental regulations. We advise that the government adopts a variety of punishments such as stepped fine system, production suspension, enforced closure or rectification of the factories, and media exposure. These measures could help completely reverse the situations including severe enterprises pollution, low illegal costs, and high law-abiding costs. Secondly, we believe that the government should reward enterprises that actively implement emission reduction technology innovation and clean production. The reward could include financial subsidies, low-interest loans, tax benefits, government preferential procurement and other policy incentives, and positive media exposure. It is necessary to increase support for the enterprises that adopt eco-friendly innovative technologies for emission reduction. This is because implementation of cutting-edge emission reduction technology is a high-risk and large investment. The long-term investment is less likely to translate into a considerable benefit in the near term. The favorable supportive measures could reduce operational risks caused by the emission reduction technology innovation to the enterprise to a certain extent. As mentioned above, with a new carrot-and-stick package of measures, the government may establish positive incentives and reverse restraint policies, namely, forming climate law enforcement atmosphere with clear rewards and punishments. By creating a law-abiding environment, we believe that companies will eventually tend to work with the interests of the government to promote the realization of legal emissions.

Formulate reasonable standard for the collection of sewage charges. At present, collection standards for sewage charges in China are generally low. The sewage charges paid by enterprises are far less than the cost of pollution control. Therefore, collection of sewage charges has not achieved policy effect. China's Regulations on the Administration of the Collection and Use of Sewage Charges has not been revised since implementation on July 1, 2003. With the continual deterioration of environmental problems, current standards for sewage discharge are far from addressing the environmental pollutions in China. Therefore, the government needs to revise the current regulations on collection of sewage charges and tighten the collection standards for sewage charges to effectively curb the number of pollutants

generated by enterprises. Central or local governments should incur higher sewage fees at a reasonable level and increase the sewage pollution management cost. The move may help mitigate the damaging effects indirectly caused by the environmental pollutions.

Enhance the ability of independent innovation and reduce enterprise cost of production and emission cut. In order to thoroughly curb environmental pollution, we must use technological progress and innovation. Innovation is a key for an enterprise's sustainable growth and the firm's competitiveness in the market. Innovation would reduce the cost of production and operation, improve the quality and quantity of products, enable enterprises to obtain the advantage in the cost, enhance the enterprise's competitiveness, and promote the enterprise's product lines. Moreover, technological innovation, clean production technologies, and process transformation can not only greatly improve the utilization efficiency of resources but also reduce water pollution. Enterprises can also develop dirt removal and filtration devices through technological innovation to reduce abatement cost. They can even turn waste into waste recycling, creating more revenue. Thus, enterprises must change their way of thinking, establish the awareness of independent innovation, and recognize the importance and urgency of independent innovation. Also, they must increase research investment by implementing more funding and recruiting more talent; for instance, by introducing high-quality talents from higher education institutions, enterprises can rebuild technological innovation teams. Purchasing environmental treatment devices domestically and internationally could help improve the research quality. Finally, enterprises should transform itself by accelerating technological innovation.

Improve the government's environmental protection work capacity and reduce pollution control costs. As mentioned above, the government's environmental governance costs reduction can improve equilibrium of environmental governance efforts. The government's environmental governance costs mainly include the engineering costs and management costs of pollution control. The following measures are needed to reduce the cost of environmental governance; for instance, actively carry out scientific and technological research, employ Internet of Things (IoT) technologies to build an integrated network of environmental pollution monitoring and reduce supervision costs, and create good government websites to engage in public participation.

6. Conclusions and Limitations

This study develops the differential game method to construct a game model of environmental pollution control between enterprises and local governments. However, most scholars adopt common static games, dynamic games, and evolutionary games. They do not take into account changes in pollution volume in the environment and the influence of one of the game's strategies on the other side of the game as the pollutants would change. The environmental pollution control process is a dynamic process. Due to the information asymmetry and the subject's limited rational participation behavior, it is difficult to achieve specific equilibrium between the participating subjects after a decision-making process. The players need to interact with each other dynamically to achieve the final equilibrium. The differential game just makes up for the shortcomings of the traditional game method. Differential game is extending the game theory to continuous time. Game participants can change their strategies in an infinitesimal time period. It fully takes into account the environmental damage and the impact of the whole game by the cumulative process of enterprises' emissions. The essence of the game is that the players make decisions in a time interval and consider a problem of intertemporal optimization. In differential games, each player has his own payment function, which depends on a deterministic dynamic system, that is, the system on which the state variable changes. In a static game, the choice of the player is called strategy, while in a differential game, the strategy is called control. In the framework of the differential game, control depends on time and system state changes, and the obtained feedback Nash equilibrium solution is Markov perfect.

Although the aim of this paper is to primarily guide future research, it may also help to get practitioners involved in the realm of environmental management and environmental performance amid ecological civilization. More theoretical and empirical researches in this area would lead to policy improvements. Specifically, research shows the following: (1) improve the weight of ecological civilization indicators in the government performance appraisal $\alpha(s)$ and formulate reasonable local government ecological civilization environmental protection performance appraisal, and reward and punishment system ϑ; (2) accelerate the establishment of reward and punishment mechanism for corporate emissions and establish reasonable standards for collecting sewage charges.

The results of the research would participate in government initiatives to address environmental issues and carry out incorporate extraorganizational policies. And the research could also highlight the potential problems after the researchers assess the government's performance in ecological progress. To sum up, polluters and local governments should recognize the benefits of environmental management and appreciate the resulting long-term performance improvements [16].

Data Availability

The numerical simulations data used to support the findings of this study were supplied by Liuwei Zhao under license and so cannot be made freely available. Requests for access to these data should be made to Liuwei Zhao (e-mail address: 136901672@qq.com).

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the National Nature Science Foundation of China (nos. 71974081 and 92046022) and Philosophy & Social Science Fund for University of Jiangsu (no. 2019SJA1051).

References

[1] H.-L. Li, X.-H. Zhu, J.-Y. Chen, and F.-T. Jiang, "Environmental regulations, environmental governance efficiency and the green transformation of China's iron and steel enterprises," *Ecological Economics*, vol. 165, Article ID 106397, 2019.

[2] K. Tevapitak and A. H. J. Helmsing, "The interaction between local governments and stakeholders in environmental management: the case of water pollution by SMEs in Thailand," *Journal of Environmental Management*, vol. 247, pp. 840–848, 2019.

[3] Y. Du, Z. Li, J. Du, N. Li, and B. Yan, "Public environmental appeal and innovation of heavy-polluting enterprises," *Journal of Cleaner Production*, vol. 222, pp. 1009–1022, 2019.

[4] P. Benhong, L. B. Yue, E. Ehsan, and W. Guo, "Dynamic evolution of ecological carrying capacity based on the ecological footprint theory: a case study of Jiangsu Province," *Ecological Indicators*, vol. 99, pp. 19–26, 2019.

[5] Y. Bian, K. Song, and J. Bai, "Market segmentation, resource misallocation and environmental pollution," *Journal of Cleaner Production*, vol. 228, pp. 376–387, 2019.

[6] G. Liu, Z. Yang, B. Chen, Y. Zhang, M. Su, and S. Ulgiati, "Prevention and control policy analysis for energy-related regional pollution management in China," *Applied Energy*, vol. 166, pp. 292–300, 2016.

[7] K. Jiang, D. You, R. Merrill, and Z. Li, "Implementation of a multi-agent environmental regulation strategy under Chinese fiscal decentralization: an evolutionary game theoretical approach," *Journal of Cleaner Production*, vol. 214, pp. 902–915, 2019.

[8] Z. Sun and W. Zhang, "Do government regulations prevent greenwashing? an evolutionary game analysis of heterogeneous enterprises," *Journal of Cleaner Production*, vol. 231, pp. 1489–1502, 2019.

[9] K. Jiang, D. You, Z. Li, and S. Shi, "A differential game approach to dynamic optimal control strategies for watershed pollution across regional boundaries under eco-compensation criterion," *Ecological Indicators*, vol. 105, pp. 229–241, 2019.

[10] X. Yunan, L. Weixin, Y. Yujie, and W. Hui, "Evolutionary game for the stakeholders in livestock pollution control based on circular economy," *Journal of Cleaner Production*, vol. 282, no. 1, Article ID 125403, 2021.

[11] J. Jiao, J. Chen, L. Li, and F. Li, "A study of local governments' and enterprises' actions in the carbon emission mechanism of subsidy or punishment based on the evolutionary game," *Chinese Journal of Management Science*, vol. 25, no. 10, pp. 140–150, 2017.

[12] F. Pan, B. Xi, and L. Wang, "Analysis on environmental regulation strategy of local government based on evolutionary game theory," *Systems Engineering-Theory & Practice*, vol. 35, no. 6, pp. 1393–1404, 2015.

[13] W. Chen and Z.-H. Hu, "Using evolutionary game theory to study governments and manufacturers' behavioral strategies under various carbon taxes and subsidies," *Journal of Cleaner Production*, vol. 201, pp. 123–141, 2018.

[14] F.-Y. Lu, "Evolutionary game analysis on environmental pollution problem," *Systems Engineering-Theory & Practice*, vol. 27, no. 9, pp. 148–152, 2007.

[15] P. Benhong, W. Yuanyuan, E. Ehsan, and W. Guo, "Behavioral game and simulation analysis of extended producer responsibility system's implementation under environmental regulations," *Environmental Science and Pollution Research*, vol. 26, no. 17, pp. 17644–17654, 2019.

[16] P. Benhong, W. Yuanyuan, Z. Sardar, W. Guo, and E. Ehsan, "Platform ecological circle for cold chain logistics enterprises: the value co-creation analysis," *Industrial Management & Data Systems*, vol. 120, no. 4, pp. 675–691, 2020.

[17] D. W. K. Yeung and L. A. Petrosyan, "A cooperative stochastic differential game of transboundary industrial pollution," *Automatica*, vol. 44, no. 6, pp. 1532–1544, 2008.

[18] G. Kossioris, M. Plexousakis, A. Xepapadeas, A. de Zeeuw, and K.-G. Mäler, "Feedback Nash equilibria for non-linear differential games in pollution control," *Journal of Economic Dynamics and Control*, vol. 32, no. 4, pp. 1312–1331, 2008.

[19] A. Yanase, "Global environment and dynamic games of environmental policy in an international duopoly," *Journal of Economics*, vol. 97, no. 2, pp. 121–140, 2009.

[20] S. Jorense and G. Zaccour, "Time consistent side payments in a dynamic game of downstream pollution," *Journal of Economic Dynamics and Control*, vol. 25, no. 12, pp. 1973–1987, 2001.

[21] S. Xu and C. Han, "Study on bain ecological compensation mechanism based on differential game theory," *Chinese Journal of Management Science*, vol. 27, no. 8, pp. 199–206, 2019.

[22] L. Musu and M. Lines, "Endogenous growth and environmental preservation," in *Environmental Economics: Proceedings of European Economic Associations at Oxford, 1993*, G. Boero and A. Silberston, Eds., St Martins Press, London, UK, 1995.

[23] L. W. Zhao and C. O. A. Otoo, "Stability and complexity of a novel three-dimensional environmental quality dynamic evolution system," *Complexity*, vol. 2019, Article ID 3941920, , 2019.

[24] Y. J. Zhang, *A Dissertation Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Management*, Huazhong University of Science and Technology, Wuhan, China, 2012.

[25] P. E. Levy and J. R. Williams, "The social context of performance appraisal: a review and framework for the future," *Journal of Management*, vol. 30, no. 6, pp. 881–905, 2004.

[26] X. Y. Wu and A. P. Pu, "On the government interest of public policies," *Journal of Yanbian Party School*, vol. 24, no. 2, pp. 37–40, 2009.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 6693730, 12 pages
https://doi.org/10.1155/2021/6693730

Research Article

An End-to-End Review-Based Aspect-Level Neural Model for Sequential Recommendation

Yupeng Liu,[1] **Yanan Zhang**,[1] **and Xiaochen Zhang**[2]

[1]*Department of Software, Harbin University of Science and Technology, Xuefu Street 52, Harbin 150001, China*
[2]*School of Materials Chemical and Engineering, Heilongjiang Institute of Technology, Harbin, China*

Correspondence should be addressed to Xiaochen Zhang; zxc161616@126.com

Received 28 November 2020; Revised 3 February 2021; Accepted 9 February 2021; Published 20 February 2021

Academic Editor: Stefania Tomasiello

Users' reviews of items contain a lot of semantic information about their preferences for items. This paper models users' long-term and short-term preferences through aspect-level reviews using a sequential neural recommendation model. Specifically, the model is devised to encode users and items with the aspect-aware representations extracted globally and locally from the user-related and item-related reviews. Given a sequence of neighbor users of a user, we design a hierarchical attention model to capture union-level preferences on sequential patterns, a pointer model to capture individual-level preferences, and a traditional attention model to balance the effects of both union-level and individual-level preferences. Finally, the long-term and short-term preferences are combined into a representation of the user and item profiles. Extensive experiments demonstrate that the model substantially outperforms many other state-of-the-art baselines substantially.

1. Introduction

In the era of information explosion, information overload is an important problem faced by users. Recommendation system came into being to solve this problem. They have evolved into a fundamental tool to help users in selecting items. Recommendation models include review-based model [1], attention-based model [2–4], aspect-based model [5], etc [6–8].

In the academic research area, a lot of work is focused on modelling long-term preferences. However, there is no appropriate model for short-term preferences from near-neighbor users. Near-neighbor users often influence the item purchase decisions made by a user. For example, a user would be highly interested in buying Apple mobile phones if his/her near-neighbor users buy them. Recently, sequential recommendations have drawn an increasing attention from both academic and industrial circles. The task is to identify an item purchased by a user by considering his or her temporal preference as a sequence.

Meanwhile, many techniques have been developed to tackle the data sparsity problem by utilizing the semantic signals provided by reviews. To model the rich semantic information in textual reviews, it is necessary to move beyond the surface-level feature representations of words in reviews. Consider the following two short sentences which contain the word "short": (1) "a short battery life," and (2) "a short loading time." Here, "short" has different sentimental information in the two sentences. Every word has different meanings in different contexts. Word embedding only considers words without considering the different meanings of each word in different contexts.

To this end, we propose an end-to-end review-based aspect-level neural model (RANM) for a sequential recommendation. Compared with the pipeline model, our end-to-end model overcomes the inherent problems of the pipeline model: time delay, parameter redundancy, and error propagation. The incorporation of reviews provides more abundant semantic information and enhances the model's expressive ability. Specifically, we embed user-related and item-related reviews into a continuous low-dimensional dense vector space by utilizing aspect-aware convolution and self-attention. Aspect-aware convolution can obtain semantic-oriented sentimental information in

reviews locally, while self-attention can obtain word association information in reviews globally. The two sides are just complementary. Then, the hierarchical attention layer is used to obtain the sequence pattern of neighbor users at the union level, and pointer mechanism is used at the individual level. Attention mechanism is used to balance the effects of the two sides. The hierarchical attention layer can obtain a multigranularity union-level neighbor sequence pattern, whereas the point mechanism can obtain an individual-level neighbor sequence pattern. The two sides are just complementary. Finally, we combine the long-term and short-term preferences to obtain a hybrid representation method for more accurate item recommendations. Overall, the key contributions of this paper towards sequential recommendation are summarized as follows:

(i) In this paper, recommendation reviews are used to precisely represent users and items. Our model can alleviate the data sparsity problem and provide a good interpretability for recommendation tasks. To the best of our knowledge, this is the first paper to harness the rich semantic information from neighbor users' reviews in the neural sequential recommendation system.

(ii) More fine-grained user and item profiles are developed. Furthermore, we use an aspect-aware convolution and a self-attention layer to identify users, their neighbors, and item representation. Meanwhile, we introduce a hierarchical attention mechanism and a review pointer to model multigranularity union-level and individual-level user preferences.

(iii) We conduct extensive experiments on a large number of datasets (Amazon and Yelp) to verify the effectiveness of the model. The importance of different components of the model and the sensitivity of the parameters are investigated. It is compared with many classic baseline systems and achieves good results on sparse datasets.

The remainder of this paper is organized as follows. In Section 2, we discuss the existing work related to our method. In Sections 3 and 4, we describe the details of RANM, model architecture, and training method. Section 5 describes our experiment setup and compares our proposed model with many state-of-the-art baselines. Finally, we conclude this paper in Section 6.

2. Related Work

2.1. Review-Based Recommendation. In order to improve the accuracy and interpretability of recommendation, there is a lot of research on utilizing reviews. Both convolution neural network (CNN) [9] and recurrent neural network (RNN) [10] have been widely adopted to extract semantic representation from reviews for rating prediction [1–4, 11–14]. DeepCoNN [13] is the first attempt to jointly model both the user and the item from reviews using neural networks. From the perspective of multimodel combination, TransNets [1] extend the DeepCoNN [13] by introducing an additional layer to represent the user-item pair. The target network assists the source network in scoring predictions. Wu et al. [11] proposed a model to exploit user-item interaction features from auxiliary users' reviews. Wu et al. [12] modelled a joint representation for a given user and item pair, which includes review-based and interaction-based feature learning. From the discovery of multiple preferences, Tay et al. [4] proposed a multipointer model that can combine multiple views of user-item interactions through review-level and word-level co-attention. Li et al. [14] considered a viewpoint of a user on an item as a semantic representation unit, which is organized into multiple logic interest units.

These methods can be classified as collaborative filtering based on interaction. From single-model view, these all are late interaction models. Our method includes not only the early interaction between neighbor users and items but also the late interaction between users and items. Our model includes a combination of the neighbor model and the user model and has a variety of preferences, such as long-term preferences and union-level and individual-level short-term preferences.

2.2. Topic-Based Recommendation. Many research studies try to extract semantic features from texts through topic model [15, 16]. CTR [17] assumes that the latent factors of items depend on the latent topic distributions of their text. TopicMF [18] links the latent topics and latent factors by using a defined transform function. McAuley et al. [19] integrated the topic method into matrix factorization through corpus likelihood regularization. Xu et al. [20] presented a model combining rating prediction and topic selection, which incorporates reviews and co-clusters of hidden user communities and item groups.

There are not only topic information but also sentimental information in reviews. Some solutions model latent factors into the framework of topic graphical model. Wang et al. [21] presented a latent aspect rating analysis model to determine the relative importance of a topical aspect. Diao et al. [22] jointly modelled ratings and review generation through aspects and their sentiments. Other solutions incorporate topical factors learned from reviews into a latent factor learning framework. Cheng et al. [23] jointly modelled the aspect-level importance and rating of reviews and item images, which were put into matrix decomposition framework for rating prediction. Shao et al. [24] matched users' and items' sentiment-aware multimodal topic models. Compared with the first kind of method, this model can be optimized both jointly and independently.

2.3. Aspect-Based Recommendation. Aspect-based recommendation systems can be divided into two main categories. The first category uses external sentiment analysis tools to extract the aspects. Zhang et al. [25] developed a multimatrix factorization model using a user-item rating matrix, a user-feature attention matrix, and an item-feature quality matrix. Chen et al. [3] learned to rank user preferences based on a phrase-level sentiment analysis across multiple categories

and further integrated this framework with matrix factorization at both the item and category levels. SLUM [26] predicts the sentiment information and then identifies the most valuable aspects of the user on that item. These works rely on the performance of the external sentiment analysis toolkit.

The second category automatically obtains the aspects of reviews by an embedded model component. Chin et al. [5] modelled the multifaceted process behind how users rate items by estimating the aspect-level user and item importance based on the neural co-attention mechanism. A^3NCF [27] defines aspects in reviews as a combination representation of topic information and embedding information. Li et al. [28] used a special CNN network to capture aspect-aware representations in reviews. Compared with the first kind of method, this explicit aspect extraction method causes error accumulation in the downstream recommendation task.

Topic-based and aspect-based recommendations are shallow latent semantic methods, but our model is a deep latent semantic method through complex hierarchical structure. In addition, our approach has a more unified model representation for joint training and inference, and the representation of latent semantics is also more diverse.

2.4. Sequential Recommendation. Sequential recommendation is an important means to use implicit feedback. Sequential pattern learning is widely verified as a critical issue in sequential recommendation [29–32]. Rendle et al. [29] presented a factorized personalized Markov chain model that subsumes both a common Markov chain and a matrix factorization model. Wang et al. [30] implemented a hierarchical representation model that can well capture both users' general tastes and sequential behaviors by involving transaction and user representations in a rating prediction. He McAuley [31] integrated similarity-based methods with Markov chains. These methods use a Markov chain to model the sequence patterns. Chen et al. [32] proposed a more flexible model to integrate collaborative filtering with a memory network.

Session/cookie-based recommendation, which does not contain user identification information, is very similar to the sequential recommendation. Hidasi et al. [33] proposed a GRU model for session recommendations. This is generally believed to be the first session recommendation based on a deep neural network (DNN). Tan et al. [34] proposed two practical techniques, namely, data augmentation and distribution shifts, to improve model performance. Jannach and Ludewig [35] proposed a heuristic near-neighbor framework for sessions that are complementary with GRU4REC [33]. Ludewig and Jannach [36] proposed an effective session-based matrix factorization, which presents the results of an in-depth performance comparison of a number of such algorithms, using a variety of datasets and evaluation measures. Our model uses an attention-based mechanism to obtain a more expressive relationship in and out of the sequence.

3. Proposed Model

In this section, we propose an end-to-end review-driven aspect-level neural recommendation model. Figure 1 illustrates the overall network architecture of the user representation network.

3.1. Model Architecture. In Table 1, let $\mathcal{U}$ and $\mathcal{I}$ represent the user and item sets, respectively. Each item $i \in \mathcal{I}$ is associated with a sequence of neighbor users arranged chronologically as $N_u = (N_u^1, \ldots, N_u^M)$, where N_u^j is a neighbor user with which the item i has interacted. We merge the user u's and the item i's review sets to form user document $D_u = (w_1, \ldots, w_{|D_u|})$ and item document $D_i = (w_1, \ldots, w_{|D_i|})$, where $|D_u|$and $|D_i|$ are the user u's and the item i 's document lengths in number of words, respectively.

Firstly, the user review D_u is transformed into an embedding matrix $\mathbf{M}_u \in R^{|D_u| \times d}$ via a neural embedding layer, where d is the dimension size of the representation vector of each word in the review. Specifically, the neural embedding layer performs a look-up operation through a shared word-embedding matrix. The word-embedding matrix can be initialized using a pretrained word vector [37, 38].

The same color is used for the same type of review and review-based vector representation in Figure 1, such as for users' reviews and review-based vectors. The same color is used for the same type of subnetwork, such as the self-attention network, CNN, and attention network.

3.2. Finding Neighbor Users. Finding neighbor users can help in determining the neighbor users in a social network. We use neighbor users' reviews to improve the user representation model, since they can provide more information for the current user. Here a neighbor user's preference is similar to that of the current user. The key here is to determine the current user's neighbor users. This paper adopts a matrix factorization method [39], a standard method to detect neighbor users. The basic idea is to decompose the rating matrix to obtain the representation of hidden variables, so as to obtain the neighbor users through a similarity function.

Specifically, the user's ratings on the item are expressed by a matrix $\mathbf{M} \in R^{|\mathcal{U}| \times |\mathcal{I}|}$, where the number of elements in the user set $\mathcal{U}$ and item set $\mathcal{I}$ is denoted by $|\mathcal{U}|$ and$|\mathcal{I}|$. Three matrices are used to approximate the rating matrix. In the learning algorithm, we use an iteration optimization method for SVD matrix factorization.

$$\min_{F \geq 0, G \geq 0, S \geq 0} \left\| \mathbf{M} - \mathbf{F}\mathbf{S}\mathbf{G}^T \right\|, \text{s.t.} \mathbf{F}^T\mathbf{F} = \mathbf{I}, \mathbf{G}^T\mathbf{G} = \mathbf{I}, \quad (1)$$

where $\mathbf{F} \in R^{|\mathcal{U}| \times K}$ represents the posterior probability of K topic clusters for each item, $\mathbf{S} \in R^{K \times K}$ represents the distribution of each topic k, and $\mathbf{G} \in R^{K \times |\mathcal{I}|}$ represents the posterior probability of k topic for each user.

We use matrix factorization to obtain the posterior probability. To infer $\mathbf{F}$, $\mathbf{S}$, and $\mathbf{G}$, the solution to the problem is obtained by the following updating rule:

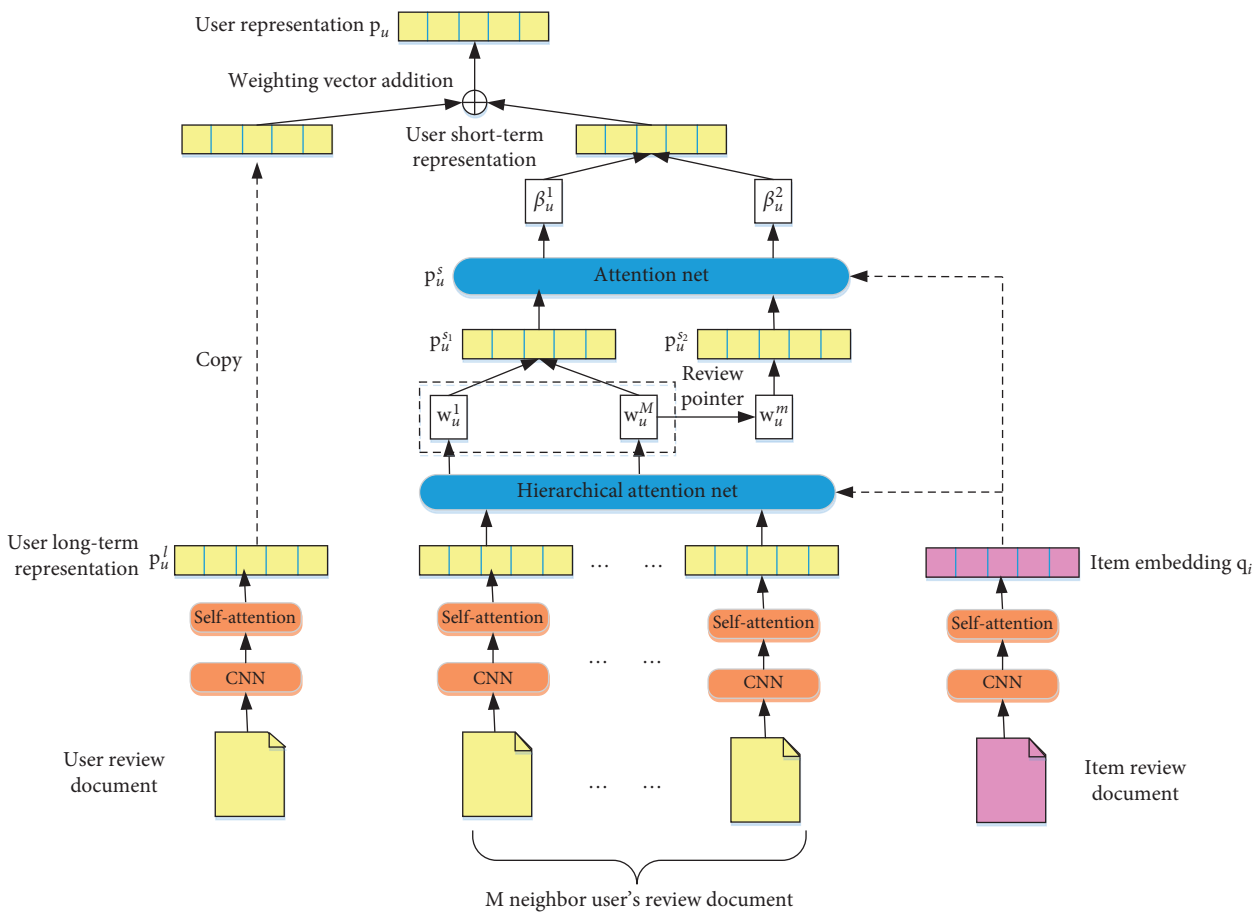


Figure 1: Network architecture of RANM.

Table 1: Notations and their definitions.

Notation	Definition
$\mathcal{U}$	User set
$\mathcal{I}$	Item set
$u \in \mathcal{U}$	User
$i \in \mathcal{I}$	Item
D_u	User u's document
D_i	Item i's document
$\mathbf{q}_i$	Item i's representation
$\mathbf{p}_u^l$	User u's long-term preference representation
$\mathbf{p}_u^{s_1}$	User u's union-level short-term preference representation
$\mathbf{p}_u^{s_2}$	User u's individual-level short-term preference representation
$\mathbf{p}_u^s$	User u's short-term preference representation
$\mathbf{q}_u^j$	User u's j-th neighbor user representation
$\mathbf{p}_u$	User u's overall preference representation
Θ_{ac}	Parameter set in CNN net
Θ_{sa}	Parameter set in self-attention net
Θ_{ha}	Parameter set in hierarchical attention net
Θ_{an}	Parameter set in attention net
r_{ui}	Rating from user u for item i

$$\mathbf{F}_{ik} \longleftarrow \mathbf{F}_{ik}\frac{\left(\mathbf{MGS}^T\right)_{ik}}{\left(\mathbf{FF}^T\mathbf{MGS}^T\right)_{ik}}, \tag{2}$$

$$\mathbf{S}_{kk'} \longleftarrow \mathbf{S}_{kk'}\frac{\left(\mathbf{F}^T\mathbf{MG}\right)_{kk'}}{\left(\mathbf{F}^T\mathbf{FSG}^T\mathbf{G}\right)_{kk'}}, \tag{3}$$

$$\mathbf{G}_{k'j} \longleftarrow \mathbf{G}_{k_1 j}\frac{\left(\mathbf{M}^T\mathbf{FS}\right)_{k'j}}{\left(\mathbf{GG}^T\mathbf{M}^T\mathbf{FS}\right)_{k'j}}. \tag{4}$$

The similarity of users u and v, given the rating matrix $\mathbf{F}$, is computed using the Pearson correlation coefficient as follows:

$$\operatorname{sim}(u, v) = \frac{\sum_{k=1}^{K}\left(\mathbf{F}_{uk} - \overline{\mathbf{F}}_u\right)\left(\mathbf{F}_{vk} - \overline{\mathbf{F}}_v\right)}{\sqrt{\sum_{k=1}^{K}\left(\mathbf{F}_{uk} - \overline{\mathbf{F}}_u\right)^2}\sqrt{\sum_{k=1}^{K}\left(\mathbf{F}_{vk} - \overline{\mathbf{F}}_v\right)^2}}, \tag{5}$$

where $\overline{\mathbf{F}}_u$ and $\overline{\mathbf{F}}_v$ denote the average weighting ratings of users u and v, respectively. The process for neighboring items is the same as for the neighbor users. Similar to equation (5), we can compute the similarity between two item vectors using a topic-item matrix $\mathbf{G}$.

3.3. Aspect-Level CNN. In this section, we describe the aspect-level convolution process of user reviews for extracting aspect-level information from reviews. Traditionally, we use the same word-embedding representation for the same word. However, the semantics or sentiment polarity of the same word can be different for different aspects. In this paper, an aspect-based transformation matrix $\mathbf{T}_a \in R^{d\times d}$ is used to transform a word-embedding matrix $\mathbf{M}_u$ into an aspect-based word-embedding matrix $\mathbf{M}_u^a = \mathbf{M}_u\mathbf{T}_\mathbf{a}$. The word embedding for K aspects can be expressed as a tensor $\mathbf{M}_u \in R^{|D_u|\times d\times K}$. Different from the traditional word-embedding representation, this paper presents the different aspects of each word.

Then, motivated by the work of Li et al. [28], we use a CNN to encode the current word's context. The tensor is similar to the multichannel feature representation of an image. In this paper, the classical image feature processing method is used to perform the convolution operation on an aspect-based feature embedding tensor. Specifically, there are n filters/kernels $\mathbf{F}^k \in R^{h\times d\times K}$, $1 \le k \le n$, where h is the height of the filter. For instance, if $n = 10$, $h = \{1, 3, 5, 7, 9\}$, then there will be two filters for each size. To obtain the nonlinear feature transformation, we use ReLU as the activation function after the convolution operation using n filters.

$$\mathbf{c}_i^k = \operatorname{ReLU}\left(\mathbf{M}_u\left[i - \frac{h}{2}: i + \frac{h}{2}\right] \circ \mathbf{F}^k + b^k\right), \tag{6}$$

where $\mathbf{M}_u[i - h/2: i + h/2]$ is the dimension h extracted from the first order of the third-order tensor, $\circ$ denotes the elementwise product of the tensor and the k-th convolution kernel $\mathbf{F}^k$, and $1 \le i \le l$is the starting position of the sliding window. Specifically, the context uses a window spanning $h/2$ words on both sides of the current word. The result is $[\mathbf{c}_1, \ldots, \mathbf{c}_l]$ ($\mathbf{c}_i \in R^n$is the context feature for the word, and n is the number of filters). Because different convolution kernels have different convolution ranges, the dimensions of the generated vectors are different. For processing convenience, different filling sizes are used for convolution kernels of different sizes. The parameter set Θ_{ac} of the aspect-level CNN layer contains $\mathbf{M}_u$, $\mathbf{T}_a$, $\mathbf{F}^k$, and b^k.

3.4. Self-Attention. In general, a pooling layer is used to refine features after the convolution layer. However, not all words are equally crucial for each user/item. Convolution considers local features, while self-attention takes global features into account. The integration of the two feature extractions can be better considered from global and local views [40]. For better generalization ability, we use the general self-attention [41].

$$\mathbf{c}_i = \sum_j \mathbf{c}_j \delta\left(\frac{\mathbf{c}_i\mathbf{W}_a\mathbf{c}_j^T}{\sqrt{n}}\right), \tag{7}$$

where $\mathbf{c}_i$ and $\mathbf{c}_j$ are the representation vectors of the current word, $\mathbf{W}_a$ are learnable parameters, and δ is a softmax function. Intuitively, the self-attention layer computes the average of all value vectors $\mathbf{c}_j$, where the weight is a matrix-weighted dot product between the query vector $\mathbf{c}_i$ and the key vector $\mathbf{c}_j$. A scaling factor n is used to avoid a large weight because of its high dimension. The parameter set Θ_{sa} of the self-attention layer contains $\mathbf{W}_a$.

The neighbor user's representation vector is encoded as follows through using max pooling. Maximum pooling implies that the strongest of these features is retained.

$$\mathbf{q}_u^j = \left[\max\left(\mathbf{c}^1\right), \ldots, \max\left(\mathbf{c}^n\right)\right], \tag{8}$$

where $\mathbf{q}_u^j$ $(1 \le j \le M)$ is the embedding for the j-th neighbor user and M is the number of neighbor users. The user's long-term preference $\mathbf{p}_u^l$ and the item representation $\mathbf{q}_i$ are extracted using the same procedure.

3.5. Hierarchical Attention Net. Hierarchical attention mechanism is used to obtain the relationship between neighbor users. Because the relationship between neighbor users is complex, hierarchical modelling is needed to obtain the relationships between different granularities. Intuitively, the order of neighbor user is pivotal for the sequential recommendation. Given a user u and his neighbor users arranged in order, each impact on the item is related to the order. Accordingly, we encode this order information by including position embedding to the user representation: $\mathbf{q}_u^j = \mathbf{q}_u^j + \mathbf{p}_u^j$ ($\mathbf{q}_u = (\mathbf{q}_u^1, \ldots, \mathbf{q}_u^{|N_u|})$, where $\mathbf{p}_u^j$ is the embedding for the j-th position.

We use the item representation $\mathbf{q}_i$ to transform the neighbor user representation $\mathbf{q}_u = (\mathbf{q}_u^1, \ldots, \mathbf{q}_u^{|N_u|})$ into a user-item representation $\mathbf{p}_u^{s_1}$as follows:

$$\mathbf{p}_u^{s_1} = \sum_i w_u^i \mathbf{q}_u^i, \tag{9}$$

where $w_u^i \in [0, 1]$is the weight of $\mathbf{q}_u^i$. The weighting function w_u^i is computed using the following equation at hop t:

$$\mathbf{o}_u^i = \tanh\left(\mathbf{W}_{q_u}\mathbf{q}_u^i + \mathbf{W}_p\mathbf{p}_u^{s_1} + \mathbf{W}_{q_i}\mathbf{q}_i\right), \tag{10}$$

$$\mathbf{w}_u^i = \frac{\exp\left(\mathbf{o}_u^i\right)}{\sum_i \exp\left(\mathbf{o}_u^i\right)}. \tag{11}$$

Note that $\mathbf{p}_u^{s_1}$ is used iteratively as a query. We define the initial value $\mathbf{p}_u^{s_1} = 1/M \sum_{i=1}^{|N_u|} \mathbf{q}_u^i$ at the first hop. Here, $\mathbf{p}_u^{s_1}$ is the sequential preference representation of neighbor users at the union level. That is, $\mathbf{p}_u^{s_1}$ is jointly encoded by all neighbor users. The parameter set Θ_{ha} of the self-attention layer contains $\mathbf{p}_u^j$, $\mathbf{W}_p$,$\mathbf{W}_{q_u}$, and $\mathbf{W}_{q_i}$.

3.6. Review Pointer. Intuitively, group-buying behavior does not represent individuality. Therefore, in addition to the union-level neighbor user representation, it also needs the individual-level neighbor user representation. For example, a user will buy a Samsung mobile phone just because most neighbor users bought a Samsung mobile phone. In this case, purchases other than Samsung are just noises, which are considered by the recommendation model. We can simply consider $\mathbf{p}_u^{s_1}$ as the final short-term preference derived from user u, but it only captures sequential patterns at the union level. To better express users' short-term preferences, we further explore the influence of neighbor users on purchase action at an individual level, that is, identify several users associated with the item.

Inspired by the pointer mechanism for review-based recommendation [4], we choose the item with the maximum attention weight as follows:

$$i = \arg\max_m \left(w_u^m\right), \tag{12}$$

$$\mathbf{p}_u^{s_2} = \mathbf{q}_u^i. \tag{13}$$

3.7. User Representation Vector. The two short-term preference representations of the user must be balanced. We utilize the attention mechanism of dot product for a weighted average of union-level and individual-level short-term preferences. The final short-term user preference is calculated as follows:

$$\mathbf{p}_u^s = \beta_u^1\mathbf{p}_u^{s_1} + \beta_u^2\mathbf{p}_u^{s_2}, \tag{14}$$

where β_u^1 and β_u^2 are the weights of union-level short-term preference $\mathbf{p}_u^{s_1}$ and individual-level short-term preference $\mathbf{p}_u^{s_2}$, respectively. The parameter set Θ_{an} of the attention layer contains $\mathbf{W}_n$, which is the matrix required to calculate attention.

The user's long-term and short-term preferences also need to be balanced. The user's current preference is formed through a linear combination of the representation vector as follows:

$$\mathbf{p}_u = (1 - \alpha)\mathbf{p}_u^l + \alpha\mathbf{p}_u^s, \tag{15}$$

where α balances the importance of the two components and is the hyperparameter. The structure of the item representation network is similar to that of the user representation network; however, it only uses the nearest neighbor items. The corresponding parameters are different.

4. Model Training

We use a factorization machine [29] to obtain the first-order and second-order features. As in collaborative filtering or matrix factorization techniques, the user and item representations are mapped into the same vector representation space $\mathscr{R}^n$. The rating $\hat{r}_{ui}$ can be computed using the inner product $p_u^T p_i$ [2, 42]. $\hat{r}_{ui}$ is a scalar, representing the user-item interaction.

$$\hat{r}_{ui} = w_0 + \sum_{i=1}^{n} w_i x_i + \sum_{i=1}^{n}\sum_{j=i+1}^{n} < \mathbf{v}_i, \mathbf{v}_j > x_i x_j. \tag{16}$$

The concatenation x of the user representation $\mathbf{p}_u$ and item representation $\mathbf{q}_i$ is passed into a factorization machine. Owing to the symmetry of the scalar product and the inner product, subscript j takes a larger value than i.

Training. This paper uses the mean squared error (MSE) as the objective function. All parameters in the component networks are trained jointly through a backpropagation training procedure. The model parameters of RANM include word embeddings, position embeddings, aspect transformation matrices, attention networks, and CNN networks.

Pretraining. We replaced the hierarchical attention layer with two feed-forward networks. One is for current user, and the other is for the interaction part between neighbor users and items. This simplified model does not consider the interaction details between neighbor users and items.

Generalization. Many research studies have found that deep neural networks tend to suffer from overfitting. Thus, we apply a L2 regularization to first-order weights, second-order weights, and biases. In addition to L2 regularization, we also use dropout [13] to further reduce overfitting. Dropout techniques can prevent co-adaptation by dropping some neural units during the training procedure [43].

5. Experiment

5.1. Experimental Setup.

Dataset: we perform our experiments on the Amazon-5cores dataset (http://jmcauley.ucsd.edu/data/amazon/). This dataset contains the product purchase histories from Amazon ranging from May 1996-July 2014. We adopt 5-core settings for these datasets. We only retain the records in Yelp in 2018 (https://www.yelp.com/dataset/challenge/) as the final dataset, which is denoted as Yelp18. For all the datasets, a standard text preprocessing method is used to process reviews. Sparsity is defined as #Rating/(#User × #Item).

For each dataset, we randomly build the training and testing sets in a ratio of 80 : 20. Moreover, 10% of the records in the training set are the development set for cross-validation. Basic statistics information of the datasets is presented in Table 2.

Metrics: here, we use the MSE as a performance metric. A statistical significance test (using * to show the result) is conducted by performing a pairwise Student's t-test.

5.1.1. Baseline Methods. To analyze our model performance more comprehensively, we use three types of baselines to conduct extensive experiments. This can not only comprehensively analyze the performance of RANM but also compare the performance of these classical baselines. We introduce the following three categories.

(a) Probabilistic matrix factorization: PMF [44].

(b) Latent topic and shallow embedding models with reviews: RBLT [45] and CMLE [46].

(c) Deep learning-based method with reviews including attention mechanism and aspect modelling: DeepCoNN [13], TransNet [1], D-Attn [2], TARMF [3], MPCN [4], ANR [5], and CARP [14].

Hyperparameter settings: We apply a grid search method for hyperparameter optimization. The final performances are reported over 5 runs. The dimension is selected from {50, 100, 200, 300}. The final dimension size of the word embedding is set to 300. The batch sizes for smaller and bigger datasets are set to be 100 and 200, respectively. For the proposed RANM, we set the number of neighbors as $M = 5$. The number of aspects is set to $K = 5$, and the embedding size is set to $d = 100$, $n = 10$, and $h = \{1, 3, 5, 7, 9\}$ for the convolution layer. The number of aspects is set as $H = 5$. The learning rate of the learning algorithm is set to 0.001 for model training, weighting factor is set as $\alpha = 0.1$ for balancing long-term and short-term preferences, and regularization coefficient is set as $\lambda = 0.0001$ to overcome overfitting.

5.2. Results and Discussion. The best configuration is used in the experiment for all baseline methods. For the deep learning method, the same TensorFlow implementation platform is used to facilitate a fair comparison. Table 3 shows the results for all 25 open datasets. Among the baselines, there is no dominating winner across all datasets. Different datasets have a certain influence on the baseline method. Overall, on most datasets, aspect-based systems (ANR and CARP) are better than review-based systems (DeepCoNN, D-Attn, TransNet, TARMF, and MPCN), where the best performance is achieved by using an attention mechanism (D-Attn, TARMF, and MPCN). The review-based system is better than the shallow semantic model (RBLT and CMLE). The shallow semantic model is better than PMF. The same conclusions can be drawn from the average performance.

Table 2: Statistics of the twenty-five datasets.

Datasets	#Users	#Items	#Ratings	Sparsity
Amazon instant video	5,130	1,685	37,126	0.0042
Apps for Android	87,271	13,209	752,937	0.00065
Automotive	2,928	1,835	20,473	0.0038
Baby	19,445	7,050	160,792	0.0012
Beauty	22,363	12,101	198,502	0.00073
Books	52, 643	91, 599	2,84, 108	0.00062
CDs and vinyl	75,258	64,443	1,097,592	0.00023
Cell phones and accessories	27,879	10,429	194,439	0.00067
Clothing shoes and jewelry	39,387	23,033	278,677	0.00031
Digital music	5,541	3,568	64,706	0.0033
Electronic	192,403	63,001	1,689,188	0.00014
Grocery and gourmet food	14,681	8,713	151,254	0.0012
Health and personal care	38,609	18,534	346,355	0.00048
Home and kitchen	66,519	28,237	551,682	0.00029
Kindle store	68,223	61,934	982,619	0.00042
Movies and TV	123,960	50,052	1,697,533	0.00027
Musical instruments	1,429	900	10,261	0.0079
Office products	4,905	2,420	53,258	0.0045
Patio, lawn, and garden	1,686	962	13,272	0.0081
Pet supplies	19,856	8,510	157,836	0.00093
Sports and outdoors	35,598	18,357	296,337	0.00045
Tools and home improvement	16,638	10,217	134,476	0.00079
Toys and games	19,412	11,924	167,597	0.00072
Video games	24,303	10,672	231,780	0.00089
Yelp18	31,831	40,841	1, 666, 869	0.0012

From all the experimental results in Table 3, the RANM method has outperformed all the baselines and has passed the hypothesis tests, demonstrating the robustness of our model. This is because a user's context better represents the short-term preferences, while other models only consider long-term preferences and do not model short-term preferences well. Overall, the experimental results demonstrate that RANM is useful in modelling sequential reviews for rating predictions. It is worthwhile to highlight that a significant improvement is gained by RANM on Yelp18 which is the sparsest dataset with the least review information in all datasets (see Table 1). Compared with the recently proposed end-to-end aspect-level models (i.e., ANR and CARP), RANM's performance improves by 8% and 18% on an average, respectively. Compared with latent topic and shallow embedding models with reviews (i.e., RBLT and CMLE), RANM's performance improved by 42% and 34% on an average. Compared with the earliest PMF model, RANM achieved an average performance improvement of up to 50%. These results show that the proposed system is sufficiently powerful.

5.3. Ablation Study. We perform an ablation study to analyze the functions of different components. The default of this discussion refers to the complete model with all components, and we compare it with its six variants:

Table 3: Performance comparison in terms of MSE.

Dataset	PMF	RBLT	CMLE	DeepCoNN	D-Attn	TransNet	TARMF	MPCN	ANR	CARP	RANM
Amazon	1.367	1.325	1.105	1.192	1.218	1.111	1.285	1.352	1.129	1.158	0.983*
Apps	1.601	1.582	1.562	1.552	1.337	1.625	1.375	1.402	1.412	1.325	**1.201***
Automotive	1.325	1.158	1.259	1.426	1.313	1.205	1.024	0.967	1.151	0.978	**0.945***
Baby	1.385	1.368	1.258	1.469	1.309	1.421	1.306	1.313	1.258	1.111	**1.009***
Beauty	1.455	1.572	1.596	1.578	1.389	1.458	1.398	1.385	1.299	1.255	**1.198***
Books	1.185	1.205	1.098	1.125	1.220	1.024	1.052	1.156	0.985	0.989	**0.942***
CDs	1.256	1.156	1.203	0.952	1.186	1.001	1.123	1.054	0.926	0.930	**0.910***
Cell phones	2.452	2.412	2.045	2.118	1.625	1.983	1.689	1.489	1.675	1.652	**1.365***
Clothing	1.528	1.569	1.452	1.425	1.384	1.356	1.386	1.124	1.255	1.023	**1.005***
Digital music	1.206	0.870	0.885	1.045	0.925	0.913	0.853	0.989	0.889	0.820	**0.789***
Electronics	1.856	1.985	1.825	1.657	1.356	1.658	1.698	1.364	1.156	1.256	**1.106***
Grocery	1.489	1.568	1.425	1.445	1.395	1.352	1.258	1.135	1.256	1.105	**1.004***
Health	1.725	1.689	1.652	1.562	1.336	1.425	1.325	1.245	1.348	1.289	**1.204***
Home	1.852	1.762	1.714	1.528	1.425	1.475	1.468	1.235	1.326	1.302	**1.210***
Kindle store	1.333	1.258	1.098	0.917	0.986	0.993	0.856	0.786	0.878	0.798	**0.724***
Movie	1.526	1.583	1.356	1.288	1.285	1.111	1.200	1.156	1.152	1.017	**1.009***
Music	1.398	0.815	0.817	1.175	1.302	0.798	0.943	0.846	1.052	0.773	**0.705***
Office	1.092	0.759	0.759	0.729	0.858	1.259	0.789	0.795	1.096	0.719	**0.706***
Patio	1.785	1.726	1.689	1.633	1.489	1.521	1.485	1.046	1.354	1.258	**1.003***
Pet supplies	1.658	1.628	1.521	1.595	1.446	1.485	1.460	1.385	1.352	1.294	**1.213***
Sport	1.428	1.596	1.324	1.301	1.275	1.215	1.115	1.134	1.103	1.059	**1.005***
Tools	1.566	0.983	1.020	1.487	1.225	1.003	1.169	1.082	1.256	0.960	**0.952***
Toys	1.396	1.425	1.305	1.236	1.106	1.112	1.213	0.985	1.093	1.007	**0.920***
Video	1.672	1.143	1.253	1.445	1.346	1.390	1.195	1.301	1.187	1.084	**0.985***
Yelp18	1.789	1.987	1.852	1.669	1.458	1.602	1.582	1.469	1.556	1.413	**1.398***
Average	1.533	1.445	1.363	1.382	1.288	1.3	1.25	1.159	1.206	1.103	**1.02***

The best and second-best results are highlighted in bold and underlined, respectively. A statistical significance test (using * to show the result) is conducted by performing a pairwise Student's t-test.

No shared projection (SP) and aspect-level projection (AP): rather than having aspect-specific transformation matrices, we constrain the model using only a single transformation matrix that is shared across all aspects of extractions. Furthermore, we remove the AP. Instead of modelling the aspect-level CNN, we do not differentiate between the aspects of each word by using a simple CNN.

No pretraining (PT): we forgo the pretraining phase (described in Section 4) for the parameters Θ_{ha} of hierarchical attention, that is, the set of parameters for the hierarchical attention layer.

No self-attention (SA): the self-attention layer is removed directly, and the global feature representation between word vectors is obtained without a self-attention mechanism.

No position embedding (PE) and sequential model (SM): on the one hand, the order of neighbors should have some impact on model performance from intuition. On the other hand, we do not use neighbor users to model the interaction with the target item. Our model can alleviate the problem of data sparsity or cold start to some extent because the SM contains rich aspect-level semantic information about long-term and short-term user preferences.

The results of the ablation study for the representative Toys, Video, Beauty, and Book datasets are shown in Table 4. According to the experiment, the removal of neighbor users results in the lowest performance, which demonstrates the importance of sequential modelling. We observe that a lack of shared AP leads to a small performance degradation. This is the smallest performance degradation of all, showing that the information captured by the shared method is not significantly different from that captured by the nonshared method. If the aspect-level project is further removed, the performance will continue to decline to the second-worst one. This shows the importance of aspect-level reviews and sequence modelling of neighbor users. It also demonstrates the effectiveness of our method in the paper.

Other factors also have a certain impact on the results. Relatively speaking, the impact is not significant. For the hierarchical attention layer, we find that the PT procedure provides a better starting optimization point for the entire optimization objective function. For the SA layer, the CNN only obtains local relevance, but not the global relevance. The SA mechanism is just a good supplement to the CNN. The no PE result suggests that temporal information in the form of a neighbor's proximity is a useful signal for a sequential recommendation.

5.4. Parameter Sensitivity.

Effects of #aspect: Figure 2(a) shows the effect of varying the number of aspects from 1 to 10. We observe that a good performance can be obtained using approximately 5 to 7 aspects. The optimal number of aspects in each dataset is slightly different because the

Table 4: Comparison of the model variants for the Toys, Video, Beauty, and Book datasets.

Architecture	Toys	Video	Beauty	Book
No SP	0.934	1.023	1.201	0.988
No AP and no SP	<u>1.035</u>	<u>1.362</u>	<u>1.385</u>	<u>1.156</u>
No PT	0.956	1.122	1.212	1.026
No SA	0.962	1.343	1.320	1.089
No PE	1.023	1.025	1.271	1.088
No SM	**1.123**	**1.401**	**1.482**	**1.293**
Default	0.920	0.985	1.198	0.942

The worst and second-worst results are highlighted in bold and underlined, respectively.

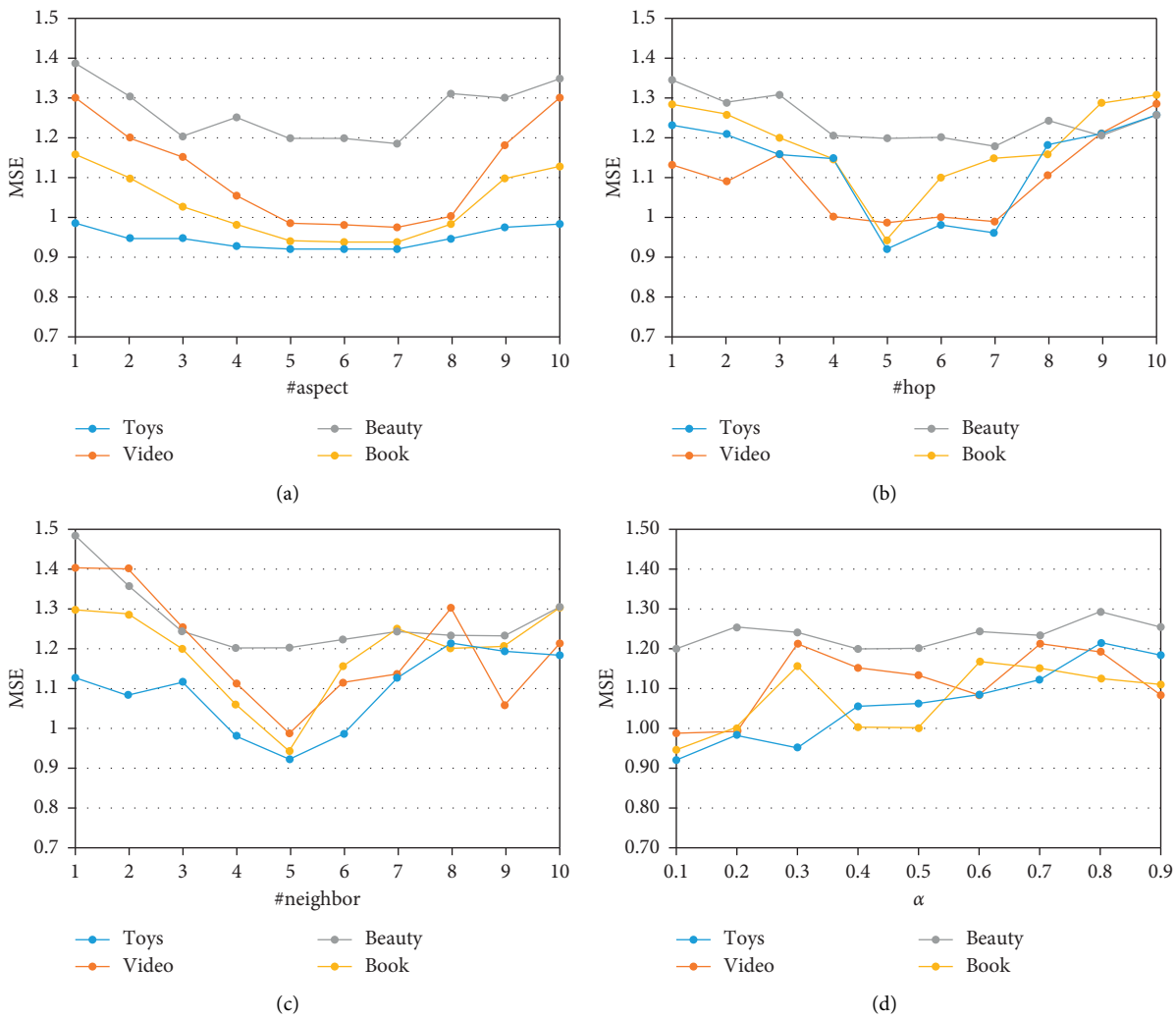


Figure 2: Parameter sensitivity of #aspect, #hop, #neighbor, and weighting parameter. (a) Effects of #aspect. (b) Effects of #hop. (c) Effects of #neighbor. (d) Effects of weighting parameter.

number of aspects in the reviews of different datasets is not the same.

Effects of #hop: we show the influence of hops of the hierarchical attention layer in Figure 2(b). Note that the model would only consider the general neighbor user's reviews when the number of hops $H = 0$. Figure 2(b) shows the effect of varying the number of hops from 1 to 10 for our model in the test datasets. This indicates that multiple hops can capture more abstract information from neighbor users' reviews as an external memory. However, too many hops lead to overfitting.

Effects of #neighbor: RANM consistently achieves a better performance at $M = 5$, as shown in Figure 2(c). When M becomes smaller or larger, the RANM performance degrades to some extent. This is reasonable because a small M produces less useful neighbor user

Table 5: Case study of two user-item pairs from office products.

User1-item1 (ink cartridge)	
User review	The HP 940XL Yellow ink cartridge came in the original sealed box & sealed wrapper. I inserted it into the printer & it has worked great.
Neighbor user review1	This was the least expensive way we could find to purchase manufacturer produced ink cartridges. It works very nicely. I just wish printer ink, in general, wasn't quite so expensive...
Neighbor user review2	Easy to install in Hewlett Packard 6500 Officejet printer, vivid color and accurate printing on HP multipurpose paper.
User2-item2 (folder)	
User review	These folders work great for my filing needs. Assorted colors of folders allow for categorization of contents. The ready tabs are easy to use with a felt tip pen, pencil or ball point.
Neighbor user review1	I like the assorted colors and love the lift up tabs that are quick and easy. However, the weight is fairly standard and not as heavy duty as I would like...
Neighbor user review2	I love these ready tab folders. So far these have been my favorite out of the hanging folders! Love them.

information and a large M would inevitably introduce much noise. Hence, we set $M = 5$ in our experiments.

Effects of weighting parameter: parameter α in equation (15) controls the importance of sequential patterns (i.e., union-level and individual-level short-term preferences). Figure 2(d) plots a performance curve with varying α values. When $\alpha = 0$, RANM degrades to a recommendation system based on reviews without sequential modelling. We observe that RANM performs much worse in this setting. An optimal performance is consistently achieved when $\alpha = 0.1$ in all the three datasets. As α becomes increasingly larger, the performance of RANM further degrades to a larger extent. These results show that short-term user preferences are a useful supplement to long-term user preference. Accordingly, we set $\alpha = 0.1$ in our experiments.

5.5. Model Interpretability. Recall that RANM encodes aspect information in the reviews. To better visualize an aspect-level review, we show the top-n phrases whose weight is the sum of the weights of the constituent words. Important phrases are highlighted in red. The weight of words in neighbor users' reviews is obtained by the weight product of equations (7) and (9). The weight of the words in the user's reviews is obtained by equation (7). We randomly sample two users and their neighbor users from office products in the Amazon dataset, as shown in Table 5. From the following two examples, we can see that the RANM model has a more fine-grained representation of users' reviews. Most of the attention to users' and neighbor users' reviews is focused on aspect and opinion words.

6. Conclusions

In this paper, we propose RANM, a novel neural recommendation model for sequential recommendation with reviews. Our model incorporates both user's long-term intrinsic preference and short-term preference to predict the user's rating of the target item. It utilizes aspect-level projections to extract aspect-level sentiment representations. The CNN and the SA layer cooperate with each other to extract global and local user features (including users to be predicted and neighbor users) and item features from the relevant review documents. A novel hierarchical attention model is proposed to capture the fine-grained alignment probability. At the same time, it uses the attention model to obtain union-level and individual-level short-term preferences. The experimental results show that RANM significantly outperforms various strong state-of-the-art methods. Experiments demonstrate the superiority of our model, especially for sparse data. Furthermore, our model can interpret the recommendation results based on reviews. In future, we will extend RANM to consider the effect of each review on the recommendation. At the same time, the user's interest of point on the item can also be modelled in a more fine-grained and adaptive learning manner.

Data Availability

We perform our experiments on the Amazon-5cores dataset (http://jmcauley.ucsd.edu/data/amazon/). This dataset contains product purchase history from Amazon ranging from May 1996 to July 2014. We adopt 5-core settings over these datasets. We only retain the records in Yelp in 2018 (https://www.yelp.com/dataset/challenge/) as the final dataset, denoted as Yelp18.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] R. Catherine and W. W. Cohen, *TransNets: Learning to Transform for Recommendation*, RecSys, London, UK, 2007.

[2] S. Seo, J. Huang, H. Yang, and Y. Liu, *Interpretable Convolutional Neural Networks with Dual Local and Global Attention for Review Rating Prediction*, RecSys, London, UK, 2017.

[3] X. Chen, Z. Qin, Y. Zhang, and T. Xu, "Learning to rank features for recommendation over multiple categories," in *Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 305–314, Pisa, Italy, July 2016.

[4] Y. Tay, A. T. Luu, and S. C. Hui, "Multi-pointer co-attention networks for recommendation," 2018, https://arxiv.org/abs/1801.09251.

[5] J. Y. Chin, K. Zhao, S. Joty, and G. Cong, "ANR: aspect-based neural recommender," in *Proceedings of the 27th ACM International Conference*, pp. 147–156, Turin, Italy, October 2018.

[6] M. Gaeta, F. Orciuoli, L. Rarità, and S. Tomasiello, "Fitted Q-iteration and functional networks for ubiquitous recommender systems," *Soft Computing*, vol. 21, no. 23, pp. 7067–7075, 2017.

[7] G. Marques, A. Respício, and A. P. Afonso, "A mobile recommendation system supporting group collaborative decision making," in *Proceedings of the 20th International Conference on Knowledge Based and Intelligent Information and Engineering Systems*, vol. 96, pp. 560–567, New York, NY, USA, September 2016.

[8] M. Hartanto and D. N. Utama, "Intelligent decision support model for recommending restaurant," *Cogent Engineering*, vol. 7, no. No. 1, Article ID 1763888, 2020.

[9] Y. Kim, "Convolutional neural networks for sentence classification," 2014, https://arxiv.org/abs/1408.5882.

[10] T. Mikolov, M. Karafi'at, L. Burget, J. Cernock'y, and S. Khudanpur, "Recurrent neural network-based language model," in *Proceedings of INTERSPEECH 2010, 11th Annual Conference of the International Speech Communication*, pp. 1045–1048, Chiba, Japan, September 2010.

[11] L. Wu, C. Quan, C. Li, and D. Ji, "PARL: let strangers speak out what you like," in *Proceedings of the 27th ACM International Conference*, pp. 147–686, Turin, Italy, October 2018.

[12] L. Wu, C. Quan, C. Li, Q. Wang, B. Zheng, and X. Luo, "A context-aware user-item representation learning for item recommendation," *ACM Transactions on Information Systems*, vol. 37, no. 2, pp. 1–29, 2019.

[13] L. Zheng, V. Noroozi, and P. S. Yu, "Joint deep modeling of users and items using reviews for recommendation," 2017, https://arxiv.org/abs/1701.04783.

[14] C. Li, C. Quan, L. Peng et al., "A capsule network for recommendation and explaining what you like and dislike," in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 275–284, Paris, France, July 2019.

[15] T. Hofmann, "Probabilistic latent semantic indexing," pp. 50–57, SIGIR, Columbia, SC, USA, 1999.

[16] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, pp. 993–1022, 2003.

[17] C. Wang and D. M. Blei, "Collaborative topic modeling for recommending scientific articles," in *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 448–456, San Diego, CA, USA, August 2011.

[18] Y. Bao, H. Fang, and J. Zhang, *Topicmf: Simultaneously Exploiting Ratings and Reviews for Recommendation*, AAAI, Palo Alto, CA, USA, 2014.

[19] J. McAuley and J. Leskovec, "Hidden factors and hidden topics: understanding rating dimensions with review test," in *Proceedings of the 7th ACM conference on Recommender systems*, vol. 8, Hong Kong, China, February 2013.

[20] Y. Xu, W. Lam, and T. Lin, "Collaborative filtering incorporating review text and co-clusters of hidden user communities and item groups," in *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*, pp. 251–260, Shanghai, China, November 2014.

[21] H. Wang, Y. Lu, and C. Zhai, "Latent aspect rating analysis on review text data: a rating regression approach," in *Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Washington, NJ, USA, July 2010.

[22] Q. Diao, M. Qiu, C.-Y. Wu, A. J. Smola, J. Jiang, and C. Wang, "Jointly modeling aspects, ratings and sentiments for movie recommendation," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, New York, NY, USA, August 2014.

[23] Z. Cheng, X. Chang, L. Zhu, R. C. Kanjirathinkal, and M. Kankanhalli, "MMALFM: explainable recommendation by leveraging reviews and images," *ACM Transactions on Information Systems (TOIS)*, vol. 37, no. 2, p. 16, 2019.

[24] X. Shao, G. Tang, and B. K. Bao, "Personalized travel recommendation based on sentiment-aware multimodal topic model," *IEEE Access*, vol. 7, pp. 113043–113052, 2019.

[25] Y. Zhang, G. Lai, M. Zhang, Y. Zhang, Y. Liu, and S. Ma, "Explicit factor models for explainable recommendation based on phrase-level sentiment analysis," in *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*, pp. 83–92, Gold Coast, Australia, September 2014.

[26] K. Bauman, B. Liu, and A. Tuzhilin, "Aspect based recommendations: recommending items with the most valuable aspects based on user reviews," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 717–725, Halifax, Canada, August 2017.

[27] Z. Cheng, Y. Ding, L. Zhu, and M. Kankanhalli, "Aspect-aware latent factor model: rating prediction with ratings and reviews," in *Proceedings of the 2018 World Wide Web Conference, International World Wide Web Conferences Steering Committee*, pp. 639–648, Hong Kong, China, May 2018.

[28] C. Li, X. Niu, X. Luo, Z. Chen, and C. Quan, "A review-driven neural model for sequential recommendation," pp. 2866–3287, 2019, https://arxiv.org/abs/1907.00590.

[29] S. Rendle, C. Freudenthaler, and L. Schmidt-Thieme, "Factorizing personalized Markov chains for next-basket recommendation," in *Proceedings of the 19th international conference on World wide web-WWW '10*, pp. 811–820, Raleigh, NC, USA, April 2010.

[30] P. Wang, J. Guo, Y. Lan, J. Xu, S. Wan, and X. Cheng, "Learning hierarchical representation model for nextbasket recommendation," in *Proceedings of the 38th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 403–412, Santiago, Chile, August 2015.

[31] R. He and J. McAuley, "Fusing similarity models with Markov chains for sparse sequential recommendation," in *Proceedings of the 2016 IEEE 16th International Conference on Data Mining*, pp. 191–200, Barcelona, Spain, July 2016.

[32] X. Chen, H. Xu, Y. Zhang et al., "Sequential recommendation with user memory networks," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pp. 108–116, Los Angeles, , CA, USA, February 2018.

[33] B. Hidasi, A. Karatzoglou, L. Baltrunas, and D. Tikk, "Session-based recommendations with recurrent neural networks," 2015, https://arxiv.org/abs/1511.06939.

[34] Y. K. Tan, X. Xu, and Y. Liu, "Improved recurrent neural networks for session-based recommendations," in *Proceedings*

of the 1st Workshop on Deep Learning for Recommender Systems, pp. 17–22, Boston, MA, USA, September 2016.

[35] D. Jannach and M. Ludewig, "When recurrent neural networks meet the neighborhood for session-based recommendation," in *Proceedings of the Eleventh ACM Conference on Recommender Systems*, Como, Italy, August 2017.

[36] M. Ludewig and D. Jannach, "Evaluation of session-based recommendation algorithms," 2018, https://arxiv.org/abs/1803.09587.

[37] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean, "Distributed representations of words and phrases and their compositionality," 2013, https://arxiv.org/abs/1310.4546.

[38] J. Pennington, R. Socher, and C. D. Manning, "GloVe: global vectors for word representation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, pp. 1532–1543, Doha, Qatar, October 2014.

[39] C. Ding, T. Li, W. Peng, and H. Park, "Orthogonal nonnegative matrix t-factorizations for clustering," in *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 126–135, Heidelberg, Germany, June 2006.

[40] W. Yu, D. Dohan, M.-T. Luong et al., "Combining local convolution with global self-attention for reading comprehension," 2018, https://arxiv.org/abs/1804.09541.

[41] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need," 2017, https://arxiv.org/abs/1706.03762.

[42] R. Fänge, U. Lidman, H. Zhang, L. Nie, X. Hu, and T.-S. Chua, "Secretion of sulfuric acid in Cassidaria echinophora Lamarck (Mollusca: mesogastropoda, marine carnivorous snail)," *Comparative Biochemistry and Physiology. A, Comparative Physiology*, vol. 53, no. 1, pp. 101–103, 1976.

[43] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[44] R. Salakhutdinov and A. Mnih, "Probabilistic matric factorization," in *NIPS'07: Proceedings of the 20th International Conference on Neural Information Processing Systems*, pp. 1257–1264, Toronto, Canada, December 2007.

[45] Y. Tan, M. Zhang, Y. Liu, and S. Ma, "Rating-boosted latent topics: understanding users and items with ratings and reviews," in *Proceedings of the 25th international conference on Machine learning-ICML'08*, pp. 2640–2646, Helsinki, Finland, July 2016.

[46] W. Zhang, Q. Yuan, J. Han, and J. Wang, "Collaborative multi-level embedding learning from reviews for rating prediction," in *Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, pp. 2986–2992, New York, NY, USA, July 2016.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 6672944, 8 pages
https://doi.org/10.1155/2021/6672944

Research Article

Analysis on Communication Influence of Official Microblogs

Nan Chen,[1] Dingguo Yu,[2,3] Yijie Zhou,[3] and Wenjuan Li[1]

[1]*Qianjiang College, Hangzhou Normal University, Hangzhou 310018, China*
[2]*College of Media Engineering, Communication University of Zhejiang, Hangzhou 310018, China*
[3]*Key Lab of Film and TV Media Technology of Zhejiang Province, Communication University of Zhejiang, Hangzhou 310018, China*

Correspondence should be addressed to Dingguo Yu; yudg@cuz.edu.cn

Received 7 December 2020; Revised 22 January 2021; Accepted 31 January 2021; Published 12 February 2021

Academic Editor: Jorge E. Macias-Diaz

As an innovative way for government affairs disclosure, official microblogs have been widely used in information release, public opinion listening, and public services, which become a model of "Internet + government affairs." The communication influence of official accounts is mainly related to the account's attention, activity, and the posted articles. Based on AHP, we design a multi-hierarchy-rank model to filter the factors and determine the most important factors related to the influence of official microblogs. This study explores the communication influence of official microblog from the perspectives of breadth, depth, and intensity; then, a three-dimensional integration of government microblogs influence calculation model for radiation, activity, and interaction is built and a monitoring and analysis system for the communication influence of government microblog is developed; finally, about 2,800 official accounts and more than 1.5 million articles in Zhejiang Province are monitored and analyzed. According to microblog verification, our proposed calculation model not only quantifies the communication influence of official microblogs but also analyzes their dynamics and timeliness. Therefore, it helps government departments to promote the healthy and orderly development of government microblogs.

1. Introduction

The new medias have entered the "Wei" era, and social medias represented by Microblog, WeChat, etc. have developed rapidly, which have become the primary platforms for the release and sharing of major information. As an innovative way for government affairs openness, the government affairs Microblog and WeChat public accounts have become more and more important, becoming the "standard configuration" for government affairs construction in many industries and regions, and being used to release information, collect opinions, listen to public opinion, and serve the public [1–3]. From the 43rd "China Internet Development Statistics Report" issued by China Network and Information Exchange Center [4], it can be obtained that, as of March 2020, the number of online government service users in China have reached 694 million, accounting for 76.8% of the overall netizens, of which the number of government microblogs certified by Sina platform have reached more than 139 thousand.

In social media, users establish connections with other users through the following function to build the topology of the social network and conduct information interactions by publishing, forwarding, and commenting information. The government affairs microblog is a bridge and link between the government and the masses. After the public becomes a fan through the following function, they can not only learn about government affairs information but also participate in government affairs interaction and obtain government services. Therefore, government affairs microblogs can make government decision-making more fair, open, and transparent, enabling people to express their opinions and suggestions without time and space restrictions and protecting people's right to speak. It collects public sentiment in time, and it is also a test of the government's administrative ability. This form can not only collect public sentiment in time but also test the government's administrative ability.

As a model of "Internet + government affairs," government affairs microblog has promoted the openness of government affairs through the application of new social

media methods, enabling the sunny government and service-oriented government affairs to be indexed and evaluated. At present, the quality of the government affairs microblog platform is uneven, and some of them are proficient; how to evaluate and measure the effectiveness of its construction is imperative. On the one hand, it is necessary to summarize and analyze the effectiveness of the government affairs microblog platform construction at all levels of the province and relevant departments to find problems; on the other hand, the "Internet + government affairs" construction model needs to be further explored, so as to promote the openness of government affairs through Internet platforms, promote the Sunshine Government, and build a service-oriented government.

Affected by the massiveness, immediacy, and global interconnectivity of Internet platform data, as well as the complexity of social media network structures and the particularity of communication methods, the traditional media communication utility metrics and evaluation methods are completely incapable of evaluating the communication influence of government affairs microblogging platforms. Therefore, the study designs a data collector to extract related data from official accounts, such as articles and comments. Then, we propose a data analysis system, which can apply naive Bayes algorithm to filter the spam articles and users' comments. Next, analytic hierarchy process (AHP) is used to mine the data and extract relevant variables according to the hierarchy, which is called multi-hierarchy-rank. For example, in the study, we propose three factors of social media, such as network topology, user interaction behavior, and information content. And in the second layer, we can exact subfactors, such as user' comments, approval, due to the three elements mentioned above. Moreover, by applying big data mining and analysis techniques, our study constructs an evaluation system and corresponding computing models of government affairs microblog communication influence from the perspectives of breadth, depth, dimension, and effect. Last but not least, a government affairs microblog communication influence monitoring and analysis system software is integrated and developed to conduct an empirical analysis of the data collection and communication effects of government affairs microblog platforms established by government agencies at various levels in Zhejiang Province.

2. Analysis of Research Status at Home and Abroad

The three basic elements of social media include network topology, user interaction behavior, and information content [5]. The topology structure reflects the intuitive characteristics of user influence from the perspective of network topology; the user interaction behaviors (likes, reposts, comments, etc.) can reflect the generation and evolution of user influence; the information content (points of interest of user, topic content, novelty, dissemination time, etc.) can further reflect the details of the generation and evolution of influence [6]. Therefore, the main research direction today is measuring and evaluating the influence of microblog users from these three aspects.

The method of measuring the influence according to topology is mainly based on the measure of node degree, which mainly includes in-degree, out-degree, and degree centrality. The degree centrality [7] is used to measure the average influence of the current node on its neighbors; the higher the centrality of the node degree, the more the objects that focus on the user, the greater the influence of the node; the centrality of the degree reflects the local characteristics of the node. Besides, local centrality and local clustering coefficient not only consider the information of the node itself but also include some information of neighbor nodes.

Chen et al. established a part local centrality that collects multilevel neighbor information to sort the influence of nodes in the network via examining the degree information of the nodes with the nearest and second nearest neighbors [8]. The global evaluation criteria of nodes mainly comprise betweenness centrality [9], closeness centrality [10], and eigenvector centrality [11].

The PageRank algorithm and its extension method obtain the user's influence ranking by taking the user's interaction in the social network as an effective link and performing statistical ranking accordingly. For example, the Topic Leader Rank algorithm proposed by Wu et al. [12] can construct user behavior network computing user impact by using content and social attributes of microblog users; the Influence Rank algorithm [13] which considers the novelty of blog text itself and the link between blog posts; microblog-Rank algorithm [14] considering the user comment relationship; a calculation model of Sina microblog user influence based on information interaction proposed by Yu et al. In addition to the following relationships, users in social media also have actions such as likes, comments, and reposts, and the strength of influence between users can be better judged by analyzing these behaviors of users. By adding user reading habits and reposting characteristics in the process of calculating user influence, Mao et al. [15] comprehensively evaluated the influence. Based on a combination of network topology and historical user behavior, Tan et al. calculated the user influence [16]. On the basis of fully researching the four basic behaviors of users' forwarding, replying, reading, and copying, the calculation method of transition probabilities inside the network is given and the random walk model Multi-Rank based on multi-relational networks is proposed to calculate users' single-topic influence and multitopic influence by Ding et al. [17].

Most social media content is short and concise, and latent dirichlet allocation (LDA) model is widely used as a topic extraction method for social media content analysis. By using LDA model to obtain the distribution of users in different topics and Gibbs to find out the strength of users' influence in different topics, Tang et al. [18] have obtained the user influence. According to Twitter's attention structure and user interest similarity, the Twitter Rank algorithm was proposed by Weng et al. [14] to calculate the user's influence on each topic; combining the social network topology with the LDM model, the social-relational topic model (SRTM) was designed by Guo et al. [19] to calculate the user's influence. In addition to the content of the information, the innovative nature of the information content and the time at which the information is released all play a role in

determining the user's influence. Song et al. [20] calculate the user's influence by incorporating the novelty of information content into the topology of social networks.

Unlike ordinary microblog users, government affairs microblogs, as bridges and bonds between the government and the masses, have special functions. Some targeted research on the influence of government affairs microblogs have been carried out by scholars at home and abroad. According to the user's forwarding behavior, Chen et al. [21] summarized the basic characteristics and determinants of government microblog communication by using cluster analysis and other methods. Aiming at user's forwarding behavior, Li et al. [22] built a relatively complete government microblog reposting scale prediction model by performing feature extraction from multiple aspects such as user characteristics, content characteristics, and time characteristics. The essential characteristics of the provincial government microblog were revealed by Sha et al. [23] from the three aspects of overall network characteristics, cohesive subgroups, and centrality. By introducing complex network structure and modular public opinion indicators, Feng et al. [24] analyzed the government information microblog information diffusion channels and influence. Luo et al. [25] obtained the relationship between the characteristics of fans and the influence of government microblog through cluster analysis; Wang et al. constructed the government affairs microblog influence evaluation system by applying rough set method [26].

3. Computational Model for Government Influence on Microblogs

The influence of social media users on communication refers to the ability by posting, reposting, commenting, and liking online behaviors and their remarks to lead public opinion in social networks that influences the spread of network topics and other users' thoughts, network behavior, and speech. The influence of government affairs microblogs is similar to that of social media opinion leaders, but it is also unique. In the study, according to the relevant connotation and model of the theory of communication influence, and the application of big data mining and analysis methods, a comprehensive evaluation index system for the influence of government affairs microblog communication is defined.

3.1. Design of Evaluation Index System. The evaluation index system for the influence of government microblog communication includes three first-level indicators, six second-level indicators, and fourteen three-level indicators that have been refined layer by layer, including radiation, activity, and interaction. The following is a one-by-one explanation and description of the indicators at all levels.

Radiation indicates the radiation range of the information published by the government's microblog account and is the criterion for judging the breadth of microblog's spread. The higher the calculated value of radiation, the larger the area spread by the content published by the account, which users can be more concerned with. The indicator is refined into the following:

(1) The number of fans: Fans can see messages posted by the government affairs microblog on their own pages. Moreover, as a "blue V" authenticated user of Sina microblog, the government affairs microblog can actively send push messages to fans in the form of private messages. Fans can also further expand the visibility of the Microblog they follow through actions such as recommendations and comments.

(2) The number of forwarded posts: Forwarding behavior is the determinant of microblog's influence expansion. When the user reads and then forwards, after this process is cycled, the number of people who read the microblog will also increase. Thus, the number of forwarded posts is directly proportional to the radiation.

In the forwarding behavior, the total number of forwarding only indicates the number of forwards. However, in the three-level indicator, the total number of forwards, the average number of forwards, and the maximum number of forwards are included in the index system, which can reflect the radiation more intuitively.

Activity indicates the activeness of government microblog in a certain period of time, including the frequency of updating microblog and the number of logins to microblog. The higher the calculated value of the activity, the more standardized the operation of the account and the ability to regularly publish blog posts, and the better the effect of disseminating information. The indicator is refined into the following:

(1) The number of posts: This indicator includes the number of microblogs published by government affairs microblogs and the number of original microblogs published by it. The number of blog posts largely reflects the activity of an account. Particularly the original microblog has a higher topic level, which can arouse the interest of users, thus leading to an increase in user comments and retweets.

(2) Active days refer to the number of times the government affairs microblog account is active, which is an important determinant of Microblog's attention. A third-level indicator is introduced under the second-level indicator, which comprises the total number of active days, the number of active days in a shorter period of time, and the number of active days in a longer period of time. According to long-term observation data, it is known that blog posts published within 7 days are the golden time period for likes and comments. After exceeding this range, the numbers of likes and comments have decreased significantly. Besides, there will be no more users to like and comment on the blog post for more than 30 days. Therefore, a longer period of time is selected as 30 days, and a shorter period of time is selected as 7 days, which can better reflect the active degree of government affairs microblog.

Interaction indicates the degree of interaction between government microblog users and other users. The main purpose of government affairs microblog is not only to disseminate government information, but also to collect and understand people's opinions to serve the people. Therefore, the degree of interaction is an important dimension for calculating the influence of government affairs microblogging. This indicator is refined into the following:

(1) Number of comments: When users generate their own thoughts or opinions on a government microblog, they will comment on it. The comments on the government affairs microblog largely reflect the public's views on certain social events and policies. The comments on Microblog can be recommented at by other users, or "@" other users to discuss. Therefore, the number of comments can reflect the level of interaction between government affairs microblog and other users. Not only is the number of comments expressed by its total number, but also the average and maximum numbers are included based on the total number, so that the number of reviews can be measured more accurately and intuitively.

(2) The number of likes, which is similar to the number of comments: When users have the same opinion on a microblog or express support for the content of the microblog but do not want to express their views, they often like the microblog. The government affairs microblog publishes some microblogs about hot events and policies. If the blog is liked, this can indicate that the information is recognized and supported by the public, which can also play a role in understanding public opinion. The number of likes is also divided into total likes, average likes, and maximum likes.

(3) The credibility from comments: When users are interested the contents released by the official account, they will comment on them. But at the same time, there will inevitably be some irrelevant information, which we call spam information. Therefore, we add another dimension to the interaction influence. We use the ratio of the number of valid comments to the total number of all comments to reflect the user's trust degree to this account.

3.2. Index Weight. The communication influence evaluation index system of the government microblog combines multiple evaluation methods, and then the experts will score to determine the weight of the indicators at all levels. By consulting experts in related fields, they will score according to the/importance of different indicators and the difficulty of obtaining data. Then the expert scoring results are processed to determine the weights of three different levels of indicators. The indicator weights are shown in Table 1.

3.3. Calculation Method of Transmission Influence. The calculation of government microblog's communication influence is based on the above evaluation indicators, including its interactive data (such as the number of likes and comments) and user information (such as the number of fans and the total number of posts).

The evaluation index system divides the communication influence into three levels of indicators and the data collected directly can only describe the current microblog characteristics. A normalized operation is used on the collected data to truly reflect the characteristics of government microblogs. The normalization method is calculated by comparing the user's current day's index value with the highest value of the index and as

$$I_{kk} = \frac{\mathrm{d}_k}{\mathrm{d}_{\max}}. \tag{1}$$

If the total comment $I11$ is taken as an example, then

$$I_{11} = \frac{\text{total number of comments on government affairs microblog}}{\text{the maximum total number of comments on all government microblogs collected}}. \tag{2}$$

The calculation of communication influence needs to be added step by step according to the indicator weight table. The index value of the lowest level indicator can be calculated directly, and the secondary indicator (such as the number of likes) is calculated by weighting the subindicator to which it belongs. By analogy, the calculated value of the first-level indicator (such as the degree of interaction) is calculated by weighting its two second-level indicators and then added. Finally, the calculation of the communication influence is obtained by adding the weighted calculations of the three dimensions. The calculation formula for the second-level indicator is shown in Formula (3):

$$I_x = I_{11} \times \lambda_a + I_{12} \times \lambda_b + I_{13} \times \lambda_c. \tag{3}$$

The calculation formula for the first-level indicator is

$$I = I_1 \times \theta_a + I_2 \times \theta_b. \tag{4}$$

The formula for calculating the communication influence is

$$\mathrm{IF} = 0.4 \times R + 0.3 \times A + 0.3 \times I. \tag{5}$$

4. Monitoring and Analysis System for the Communication Influence of Official Microblogy

4.1. Structural Design of System Functions. According to the designed calculation model of communication influence of the government microblog, we have developed a

Table 1: Temperature and wildlife count in the three areas covered by the study.

First-level index	Weights	Second-level index	Weights	Third-level index	Weights
Radiation R	$\beta_1 = 0.4$	Number of forwards R_1	$\theta_1 = 0.6$	Total number of forwards R_{11}	$\lambda_1 = 0.6$
				Average number of forwards R_{12}	$\lambda_2 = 0.3$
				Maximum number of forwards R_{13}	$\lambda_3 = 0.1$
		Number of fans R_2	$\theta_2 = 0.4$	None	None
Activity A	$\beta_2 = 0.3$	Number of posts A_1	$\theta_3 = 0.7$	Total number of posts A_{11}	$\lambda_4 = 0.5$
				Number of originals A_{12}	$\lambda_5 = 0.5$
		Activity A_2	$\theta_4 = 0.3$	Total activity A_{21}	$\lambda_6 = 0.5$
				Total activity in a long period A_{22}	$\lambda_7 = 0.2$
				Total activity in a short period A_{23}	$\lambda_8 = 0.3$
Interaction I	$\beta_3 = 0.3$	Number of comments I_1	$\theta_5 = 0.5$	Total number of comments I_{11}	$\lambda_9 = 0.5$
				Average number of comments I_{12}	$\lambda_{10} = 0.3$
				Maximum number of comments I_{13}	$\lambda_{11} = 0.2$
		Number of likes I_2	$\theta_6 = 0.5$	Total number of likes I_{21}	$\lambda_{12} = 0.5$
				Average number of likes I_{22}	$\lambda_{13} = 0.3$
				Maximum number of likes I_{23}	$\lambda_{14} = 0.2$
		Credibility I_3	$\theta_7 = 0.2$	Ratio of the number of valid comments to total number of comments I_{31}	$\lambda_{15} = 1$

communication influence monitoring and analysis system software of official microblogs, comprising three major modules of microblog data directional collection, analysis engine, and visual analysis, and the software structure is shown in Figure 1.

The microblog data directional collection is mainly used to collect the data of the microblog, which can automatically collect microblog data of monitored accounts according to the configuration, comprising the content of each blog post and the corresponding statistics of the blog posts, likes, reposts, and comments. In addition, it automatically updates the statistics of blog posts on a regular basis. The data directional collection module is mainly composed of two parts of the acquisition management module and the microblog crawler. The collection management module can centrally manage and control the numerous microblog crawlers deployed in a distributed manner; the microblog crawler can realize the automatic collection of microblog web page information and can intelligently parse the web page content, automatically deduplicate, denoise, and extract the specified information content, and finally convert it into structured data.

The analysis engine can automatically perform statistical analysis on the collected microblog data and calculate the monitored influence of the government microblog account based on the defined government influence microblog's propagation influence calculation model. The visual analysis module presents the calculation results with intuitive images such as graphs and tables.

4.2. Experiment and Result Analysis. First, more than 2,600 official microblog accounts in Zhejiang Province are collected, involving public security, culture, tourism, transportation, taxation, party and government agencies, social teams, and other industries. Then, the monitoring and analysis system for the communication influence of government microblogs is used to monitor and analyze them. The system automatically collected the likes, forwards, and comments of each article published by these accounts in the past 3 months. As of December 30, 2019, the system has collected a total of 1.03 million microblog articles and about 3.56 million comments. Based on this data, the influence of official microblog account monitored by the system is calculated according to the government influence microblog communication influence calculation model proposed in the study.

The information of Top 30 government affairs microblog accounts obtained by the model is shown in Table 2.

From the data obtained from the analysis system, it can be seen that the government microblogs of government agencies, tourism, and public security have a stronger influence on social media users. The government department microblogs, which are officially set by the People's News Government Office, rank first in the influence of Zhejiang government microblogs. The articles published by the official microblogs of government agencies at all levels cover a wide range of fields, containing culture, scenery, education, security, and people's livelihood. Thus, most users will use this type of government microblogs to learn what has happened at the present moment. Therefore, the microblogs of the government agency sector rank first. Next, the influence of travel microblogs ranks second. Obviously, Zhejiang has always been famous for its rich tourism resources, beautiful scenery, and complete supporting facilities. Therefore, tourists will learn about tourism projects and activities of Zhejiang Province from the microblogs' articles of tourism departments. What is more, the microblogs' influence of public security ranks third. Users can obtain safety information and knowledge by such microblogs and try their best to prevent accidents.

In addition, the other influential factors of government affair microblogs' communication influence in Zhejiang Province are shown in Figure 2. It obviously can be seen that microblog's fan number, article number, and interaction

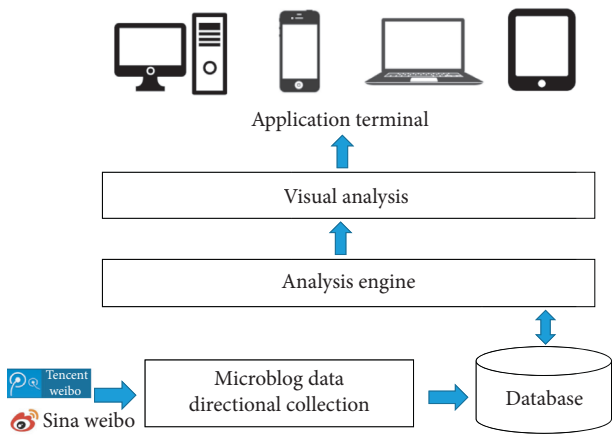


Figure 1: The structure of influence monitoring and analysis system.

Table 2: Communication influence ranking table.

Influence ranking	Names of official microblogs	Radiation index	Activity index	Interaction index	Influence index
1	Hangzhou publicity	58707	54880	23095	48251
2	Yuhang publicity	59621	38068	27480	48043
3	Zhejiang travel	58784	31920	28460	47088
4	Zhejiang fire	37734	30061	24168	33016
5	Hangzhou police	33675	43601	26124	32728
6	Zhuji education	38379	17242	22865	31521
7	Hangzhou travel	32525	40410	23576	30944
8	Zhejiang committee	25989	47662	29668	29586
9	Zhejiang publicity	27341	42840	27392	29190
10	Peace wenzhou	24362	26036	38835	28626
11	Ningbo police	22638	67075	24911	28534
12	Ningbo publicity	20189	59135	28554	27146
13	Jinhua police	18318	37771	31199	24237
14	Chunan publicity	19100	44447	24486	23611
15	Peace jianggan	14902	21300	42578	23433
16	Jiaxing fire	23238	22231	22755	22983
17	Bingjiang publicity	18732	21173	31545	22620
18	Jianggan publicity	19169	36183	23052	22272
19	Shangcheng publicity	21625	30005	19799	22103
20	Zhejiang police	17683	34887	25303	21858
21	Hanzhong urban management	14950	39697	28278	21621
22	Wenzhou traffic	14553	43372	26931	21481
23	Peace west lake	19404	17440	26016	21029
24	Huangyan publicity	16850	12931	31121	20395
25	Shangcheng petition	7754	15755	29003	20186
26	Haining publicity	14219	29224	28984	20142
27	Zhejiang law	21272	9841	20380	19669
28	Yiwu fire	7025	20310	26254	19616
29	Qiandao lake travel	15845	33727	20344	19224
30	Zhejiang weather	12893	38300	23701	18935

number make great contribution to communication influence of microblogs. Those with a large number of fans, articles, and interactions (likes, comments, and forwards) commonly have great influence on social media users. To be fair, compared with the three factors, the article number of microblogs has a weaker impact on communication influence of microblogs.

All in all, these microblogs of strong communication influence have the characteristics of high quality contents, large numbers of fans, fast update speed, frequent

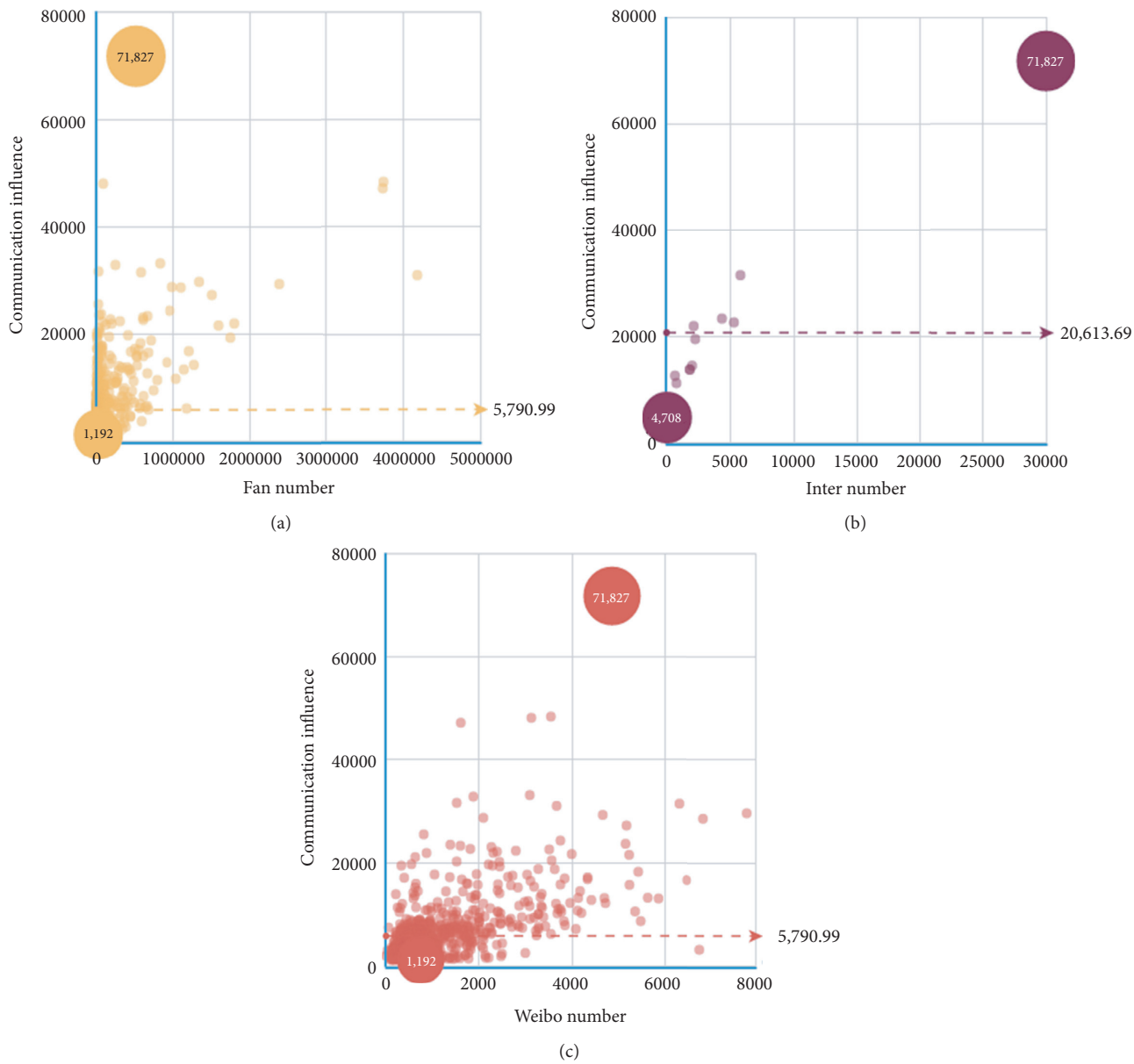


Figure 2: Influential factors of government affairs microblogs' communication influence.

interaction, etc., which attract the attention of the public, thus having a higher impact.

5. Conclusions

The government affairs microblog with a wide spread and large influence will have a profound impact on social issues, citizen deliberations, and popularization of policies. However, the spread and operation of various government microblogs are different, and homogeneous government microblogs still exist. In response to this situation, to better perform quantitative analysis and evaluate the effect and influence of government affairs microblogs, a set of government affairs microblog communication effect evaluation system and a monitoring and analysis system of micro government communication influence is designed in this study, thus analyzing and judging the government affairs microblog's communication effect and influence through a scientific and effective method. According to the practical data analysis results, the trinity model of governmental microblog communication impact based on radioactivity, activity, and interaction designed in this study can better analyze and evaluate the communication influence of government affairs microblog; besides, it can also reflect the dynamic and timeliness of communication influence of government affairs microblog. Moreover, the results indicate that the number of fans and posts is an important factor influencing the influence of microblog accounts, but it is not a decisive factor.

Data Availability

The data in this paper are derived from Sina Web account.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the MOE (Ministry of Education in China) Project of Humanities and Social Sciences (no. 18YJA860020), the National Natural Science Foundation of China (no. 61702151), and the Key Research and Development Program of Zhejiang Province, China (no. 2019C03138).

References

[1] J. Chen and Y. Qin-Jian, "A review of administrative microblogging study in China and abroad," *Information Science*, vol. 32, no. 6, pp. 156–161, 2014, in Chinese.

[2] J. Yu, Z. Xie, and X. Ma, "Development context and future prospect: ten years of research on new media of Chinese government affairs," *Academic Library and Information Science*, vol. 38, no. 05, pp. 122–129, 2020, in Chinese.

[3] L. Zhao, "The development and communication of new media of government affairs," *Reporter Observed*, vol. 24, pp. 22-23, 2020, in Chinese.

[4] CNNIC, *45th China Internet Development Statistics Report*, China Internet Network Information Central, Beijing, China, 2020, in Chinese.

[5] Z. M. Han, Y. Chen, W. Liu, B. H. Yuan, M. Q. Li, and D. G. Duan, "Research on node influence analysis in social networks," *Ruan Jian Xue Bao/Journal of Software*, vol. 28, no. 1, pp. 84–104, 2017, in Chinese.

[6] C. X. Wang, X. H. Guan, T. Qin, and W. Li, "Who are active? An in-depth measurement on user activity characteristics in sina microblogging," in *Proceedings of the GLOBECOM. Piscataway*, pp. 2083–2088, IEEE, Anaheim, CA, USA, December 2012.

[7] L. C. Freeman, "Centrality in social networks conceptual clarification," *Social Networks*, vol. 1, no. 3, pp. 215–239, 1979.

[8] D. Chen, L. Lü, M.-S. Shang, Y.-C. Zhang, and T. Zhou, "Identifying influential nodes in complex networks," *Physica A: Statistical Mechanics and Its Applications*, vol. 391, no. 4, pp. 1777–1787, 2012.

[9] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, vol. 40, no. 1, pp. 35–41, 1977.

[10] G. Sabidussi, "The centrality index of a graph," *Psychometrika*, vol. 31, no. 4, pp. 581–603, 1966.

[11] K. Stephenson and M. Zelen, "Rethinking centrality: methods and examples," *Social Networks*, vol. 11, no. 1, pp. 1–37, 1989.

[12] W. U. Xian-Hui, H. Zhang, X.-J. Zhao et al., "Mining algorithm of microblogging opinion leaders based on user-behavior network," *Application Research of Computers*, vol. 32, no. 4, 2015, in Chinese.

[13] S. Xiao-Dan, C. H. I. Yun, K. Hino et al., "Identifying opinion le aders in the blogosphere," in *Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management, 2007*, pp. 971–974, ACM Press, New York, NY, USA, January 2007.

[14] L. Yan, L. Hua-Xian, L. Xue-Qiao et al., "Hot topic propagation model and opinion leader identifying model in microblog network," *Abstract and Applied Analysis*, vol. 36, no. 2, pp. 360–367, 2013.

[15] J. X. Mao, Y. Q. Liu, H. M. Zhang, and S. P. Ma, "Social influence analysis for micro-blog user based on user behavior," *Chinese Journal of Computers*, vol. 37, no. 4, pp. 791–800, 2014.

[16] Z. M. Han, Y. Chen, W. Liu, B. H. Yuan, M. Q. Li, and D. G. Duan, "Research on node influence analysis in social networks," *Ruan Jian Xue Bao/Journal of Software*, vol. 28, no. 1, pp. 84–104, 2017, in Chinese.

[17] Z. Ding, B. Zhou, Y. Jia et al., "Topic influence analysis based on the multi-relational network in microblogs," *Journal of Computer Research and Development*, vol. 50, no. 10, pp. 2155–2175, 2013.

[18] J. Tang, J. Sun, C. Wang, C. Wang, and Z. Yang, "Social influence analysis in large-scale networks," in *Proceedings of the 15th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2009)*, J. Elder and F. S. Fogelman, Eds., ACM Press, New York, NY, USA, pp. 807–816, June 2009.

[19] W. Guo, S. Wu, L. Wang, and T. Tan, "Social-Relational topic model for social networks," in *Proceedings of the 24th ACM Int'l Conf. on Information and Knowledge Management (CIKM 2012)*, pp. 1731–1734, ACM Press, New York, NY, USA, October 2015.

[20] X. Song, Y. Chi, H. Koji et al., "Identification opinion leader in the blogosphere," in *Proceedings of the 16th ACM Conf on Information and Knowledge Management*, pp. 971–974, ACM, New York, NY, USA, January 2007.

[21] R. Chen and Y. Liu, "Study on model of government micro blogs information transmission based on forward behavior," *E-goverment*, vol. 7, pp. 108–117, 2017, in Chinese.

[22] Q. Li, J. Jiang, Li Ying, and Y. Liu, "The retweeting scale classification prediction of government microblogs in China," *Journal of Intelligence*, vol. 37, no. 1, pp. 95–99, 2018, in Chinese.

[23] S. Zhongyong and Y. Su, "Social network analysis of provincial government micro blogs in China," *Jinan Journal (Philosophy and Social Sciences)*, vol. 6, pp. 125–132, 2018, in Chinese.

[24] H. Feng, C. Lei, C. Yang, and M. Yuan, "Research on the influence of government affairs microblog based on modularity of complex network," *Journal of Changzhou University (Social Science Edition)*, vol. 17, no. 6, pp. 37–42, 2016, in Chinese.

[25] Y. Luo, G. Hu, and M. Lu, "Study on government affairs micro-blog influence and fans' characteristic relation," *E-goverment*, vol. 12, pp. 82–89, 2017, in Chinese.

[26] S. Wang, Y. Ge, and Q. Zhao, "Evalution method and empirical study of government micro-blog influence based on rough set theory," *Mathesmatics in Practice and Theory*, vol. 48, no. 22, pp. 76–84, 2018, in Chinese.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 6681932, 10 pages
https://doi.org/10.1155/2021/6681932

Research Article

On New Modifications of Some Perturbation Procedures

A. I. Ismail [1,2]

[1] *Mechanical Engineering Department, College of Engineering and Islamic Architecture, Umm Al-Qura University, Makkah, Saudi Arabia*
[2] *Mathematics Department, Faculty of Science, Tanta University, Tanta, P.O. Box 31527, Egypt*

Correspondence should be addressed to A. I. Ismail; aiismail@uqu.edu.sa

Received 29 October 2020; Revised 25 December 2020; Accepted 28 January 2021; Published 9 February 2021

Academic Editor: Jorge E. Macias-Diaz

In this paper, we present new modifications for some perturbation procedures used in mathematics, physics, astronomy, and engineering. These modifications will help us to solve the previous problems in different sciences under new conditions. As problems, we have, for example, the rotary rigid body problem, the gyroscopic problem, the pendulum motion problem, and other ones. These problems will be solved in a new manner different from the previous treatments. We solve some of the previous problems in the presence of new conditions, new analysis, and new domains. We let complementary conditions of such studied previously. We solve these problems by applying the large parameter technique used by assuming a large parameter which inversely proportional to a small quantity. For example, in rigid body dynamic problems, we take such quantity to be one of the components of the angular velocity vector in the initial instant of the rotary body about a fixed point. The domain of our solutions will be depending on the choice of a large parameter. The problem of slow (weak) oscillations is considered. So, we obtain slow motions of the bodies instead of fast motions and find the solutions of the problem in present new conditions on both of center of gravity, moments of inertia, and the angular velocity vector or one of these parameters of the body. This study is important for aerospace engineering, gyroscopic motions, satellite motion which has the correspondence of inertia moments, antennas, and navigations.

1. Introduction

In [1], the variation of parameters method is applied to solve the general nonhomogenous linear differential equation of the second order. The author solved the homogenous equation and constructed a particular solution depending on its solutions. He got the first and second differentiation of the particular solutions, substituted them into the considered equation, and performed some investigated steps, and he found the required general solution.

We note that this is an exact solution for a linear differential equation and not an approximated one and so is not the small parameter technique or the large parameter one useful for this problem. Such techniques are useful only for approximated problems consisting of nonlinear differential equations containing a small parameter [2] or a large one [3, 4]. In the following sections, we modify the well-known previous method in terms of the large parameter for solving some systems for some perturbed problems [5, 6] containing quasi-linear systems. This modification gives us the chance to study the considered problems in new conditions and new domains of solutions.

2. The Method of Averaging

We consider the following examples for illustrating this method, and we will find the periodic solutions of each case:

2.1. Van Der Pol's Technique. This technique is considered in previous works for finding the approximated periodic solutions as power series in terms of a small parameter defined for each problem. In our work, we aim to find new periodic solutions in terms of a large parameter instead of the small one of the equation

$$\frac{d^2u}{dt^2} + \omega_0^2 u = \varepsilon^{-1}\left[\left(1-u^2\right)\frac{du}{dt} + K\lambda \cos \lambda t\right], \qquad (1)$$

where ε is imposed to be large and $\lambda = O(\varepsilon^{-1})$ is the movement frequency which is assumed to be small to be different from the natural frequency ω_0. By taking the solution to (1), we have assumed the following form:

$$u(t) = a_1(t)\cos \lambda t + a_2(t)\sin \lambda t, \tag{2}$$

where $a_1(t)$ and $a_2(t)$ are taken to be slow functions that vary with time as well as $(\mathrm{d}a_i/\mathrm{d}t) = 0(\varepsilon^{-1})$ and $(d^2a_i/dt^2) = 0(\varepsilon^{-2})$. By differentiating (2) twice, we obtain

$$\ddot{u} = -\lambda^2 a_1 \cos \lambda t - \lambda^2 a_2 \sin \lambda t - 2\dot{a}_1\lambda \sin \lambda t + 2\dot{a}_2\lambda \cos \lambda t + \ddot{a}_1 \cos \lambda t + \ddot{a}_2 \sin \lambda t, \qquad \left(\dot{u} \equiv \frac{\mathrm{d}u}{\mathrm{d}t}\right). \tag{3}$$

Using equations (1)–(3), we obtain

$$\begin{aligned} 2\dot{a}_1 + a_2\left(\lambda - \frac{\omega_0^2}{\lambda}\right) - \varepsilon^{-1}a_1(1-\rho) = 0, \\ 2\dot{a}_2 - a_1\left(\lambda - \frac{\omega_0^2}{\lambda}\right) - \varepsilon^{-1}a_2(1-\rho) = \varepsilon^{-1}K, \end{aligned} \tag{4}$$

$$\rho = 0.25a^2 = 0.25\left(a_1^2 + a_2^2\right). \tag{5}$$

Periodic solutions of (1) match solutions:

$$2\sigma a_{20} - a_{10}(1-\rho_0) = 0, \tag{6}$$

$$-2\sigma a_{10} - a_{20}(1-\rho_0) = K, \tag{7}$$

where $\sigma = \lambda - \varepsilon\omega_0$.

By adding squares (6) and (7) and using (5), we obtain the solution of the frequency equation:

$$\rho_0\left[4\sigma^2 + (1-\rho_0)^2\right] = 0.25K^2. \tag{8}$$

2.2. The Krylov–Bogoliubov Technique. This technique considered through the weak nonlinear equation of the second order in terms of the small parameter was defined in the problem [7–11]. Here, we consider applying the large parameter ε. Let us consider the following equation:

$$\frac{\mathrm{d}^2u}{\mathrm{d}t^2} + \omega_0^2 u = \varepsilon^{-1} f\left(u, \frac{\mathrm{d}u}{\mathrm{d}t}\right). \tag{9}$$

At $\varepsilon \longrightarrow \infty$, the solution of (9) can be written as follows:

$$u = a\cos(\omega_0 t + \theta), \tag{10}$$

where a and θ are constants. To determine the approximated solution of (9) for the large parameter ε, Krylov and Bogoliubov hypothesized that the solution is still given by (10) but with the time change a and θ, and it is subjected to the condition:

$$\begin{aligned} \frac{\mathrm{d}u}{\mathrm{d}t} &= -a\omega_0 \sin \phi, \\ \phi &= \omega_0 t + \theta. \end{aligned} \tag{11}$$

By differentiation (10) w. r. t.t, we obtain

$$\frac{\mathrm{d}u}{\mathrm{d}t} = -a\omega_0 \sin \phi + \frac{\mathrm{d}a}{\mathrm{d}t}\cos \phi - a\frac{\mathrm{d}\theta}{\mathrm{d}t}\sin \phi. \tag{12}$$

So,

$$\frac{\mathrm{d}a}{\mathrm{d}t}\cos \phi - a\frac{\mathrm{d}\theta}{\mathrm{d}t}\sin \phi = 0. \tag{13}$$

By differentiation (11) w. r. t.t, we obtain

$$\frac{\mathrm{d}^2u}{\mathrm{d}t^2} = -a\omega_0^2 \cos \phi - \omega_0\frac{\mathrm{d}a}{\mathrm{d}t}\sin \phi - a\omega_0\frac{\mathrm{d}\theta}{\mathrm{d}t}\cos \phi. \tag{14}$$

Substituting this expression into (9) and using (10), we obtain

$$\begin{aligned} \omega_0\frac{\mathrm{d}a}{\mathrm{d}t}\sin \phi + a\omega_0\frac{\mathrm{d}\theta}{\mathrm{d}t}\cos \phi = -\varepsilon^{-1}f, \\ f = f\left[a\cos \phi, -a\omega_0 \sin \phi\right]. \end{aligned} \tag{15}$$

Solving (13) and (15) for $(\mathrm{d}a/\mathrm{d}t)$ and $(\mathrm{d}\theta/\mathrm{d}t)$, we obtain that

$$\begin{aligned} \omega_0\frac{\mathrm{d}a}{\mathrm{d}t} &= -\varepsilon^{-1}\sin \phi\, f, \\ a\omega_0\frac{\mathrm{d}\theta}{\mathrm{d}t} &= -\varepsilon^{-1}\cos \phi\, f. \end{aligned} \tag{16}$$

By integrating (16) into the period $[t, t+T]$, where a and θ can be taken as constants on the right-hand side, we obtain

$$2\omega_0\frac{\mathrm{d}a}{\mathrm{d}t} = -\varepsilon^{-1}f_1(a),$$
$$2a\omega_0\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\varepsilon^{-1}g_1(a), \tag{17}$$

where

$$\pi f_1(a) = \int_0^{2\pi} \sin\phi f\,\mathrm{d}\phi,$$
$$\pi g_1(a) = \int_0^{2\pi} \cos\phi f\,\mathrm{d}\phi. \tag{18}$$

We also note that f_1 and g_1 are coefficients of the Fourier series expansion of the function f.

As an example, let us consider the Duffing equation.

In this example, we modify the Duffing equation in terms of the large parameter to be

$$\frac{\mathrm{d}^2u}{\mathrm{d}t^2} + \omega_o^2 u = -\varepsilon^{-1}u^3, \tag{19}$$

where

$$f(u,\dot{u}) = -u^3. \tag{20}$$

So,

$$f_1(a) = 0,$$
$$g_1(a) = -0.75a^3. \tag{21}$$

Thus, from (17), it is obtained

$$\omega_o\theta = 0.375\varepsilon^{-1}a^2t + \omega_o\theta_o, \quad a = \text{const}. \tag{22}$$

Therefore, the first approximation is

$$u = a\cos\omega_o\left[1 + \left(0.375\varepsilon^{-1}a^2/\omega_o^2\right)\right]t + o\left(\varepsilon^{-2}\right). \tag{23}$$

For the second example, consider the Van der Pol oscillator:

$$\frac{\mathrm{d}^2u}{\mathrm{d}t^2} + \omega_o^2 u = f(u,\dot{u}), \tag{24}$$

where

$$f(u,\dot{u}) = \left(1-u^2\right)\frac{\mathrm{d}u}{\mathrm{d}t}. \tag{25}$$

In this case,

$$f_1 = -\omega_o a\left(1 - 0.25a^2\right),$$
$$g_1 = 0. \tag{26}$$

Hence and from (17), it follows that $(\mathrm{d}a/\mathrm{d}t)$ is obtained in terms of the large parameter, so

$$\frac{\mathrm{d}a}{\mathrm{d}t} = 0.5\varepsilon^{-1}a\left(1 - 0.25a^2\right),$$
$$\theta = \theta_o = \text{const}. \tag{27}$$

By integrating the first equation into (27), it becomes clear that

$$a^2 = \frac{4}{\left[1 + \left((4/a_o) - 1\right)e^{\varepsilon t}\right]}. \tag{28}$$

2.3. The Generalized Method of Averaging. This method was applied for obtaining the periodic solutions of the differential equations applying the small parameter technique which was defined for each case of the motion [12–14]. In our work, we use the large parameter method for obtaining new solutions dependent on the large parameter ε. Let us consider (10) and (11) as the conversion from u and $(\mathrm{d}u/\mathrm{d}t)$ to a and ϕ in terms of the large parameter, so

$$\omega_o\frac{\mathrm{d}a}{\mathrm{d}t} = -\varepsilon^{-1}\sin\phi f,$$
$$a\omega_o\frac{\mathrm{d}\phi}{\mathrm{d}t} = a\omega_o^2 - \varepsilon^{-1}\cos\phi f, \tag{29}$$

where ϕ is called the rapid rotation phase. Let us take the conversions from (a,ϕ) to $(\overline{a},\overline{\phi})$ such that

$$a = \overline{a} + \varepsilon^{-1}a_1(\overline{a},\overline{\phi}) + \varepsilon^{-2}a_2(\overline{a},\overline{\phi}) + \ldots,$$
$$\phi = \overline{\phi} + \varepsilon^{-1}\phi_1(\overline{a},\overline{\phi}) + \varepsilon^{-2}\phi_2(\overline{a},\overline{\phi}) + \ldots. \tag{30}$$

Using (30), we find that system (29) takes the following form:

$$\frac{\mathrm{d}\overline{a}}{\mathrm{d}t} = \varepsilon^{-1}A_1(\overline{a}) + \varepsilon^{-2}A_2(\overline{a}) + \ldots,$$
$$\frac{\mathrm{d}\overline{\phi}}{\mathrm{d}t} = \omega_o + \varepsilon^{-1}B_1(\overline{a}) + \varepsilon^{-2}B_2(\overline{a}) + \ldots, \tag{31}$$

where A_i and B_i do not depend on $\overline{\phi}$, for each $i = 1, 2$.

Using (29)–(31), we obtain

$$\omega_o\frac{\partial a_n}{\partial\overline{\phi}} + A_n = F_n(\overline{a},\overline{\phi}),$$
$$\omega_o\frac{\partial\phi_n}{\partial\overline{\phi}} + B_n = G_n(\overline{a},\overline{\phi}). \tag{32}$$

Generally, F_n and G_n contain short periodic terms (referred to as superscript s) and long periodic terms (referred to by superscriptℓ).

We choose A_n and B_n such that it is equal to the long periodic terms, so

$$A_n = F_n^{\ell},$$
$$B_n = G_n^{\ell}, \tag{33}$$

then

$$\omega_o \frac{\partial a_n}{\partial \overline{\phi}} = F_n^s,$$
$$\omega_o \frac{\partial \phi_n}{\partial \overline{\phi}} = G_n^s, \tag{34}$$

which can be solved, respectively, in a_n and ϕ_n.

As an example, consider the Van der Pol oscillator (24) in which

$$f(u, \dot{u}) = \left(1 - u^2\right)\dot{u}, \quad \omega_o = 1. \tag{35}$$

In this case, equation (29) in terms of the large parameter becomes:

$$\frac{\mathrm{d}a}{\mathrm{d}t} = 0.125\varepsilon^{-1}\left[a\left(4 - a^2\right) - 4a \cos 2\phi + a^3 \cos 4\phi\right],$$
$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = 1 + 0.125\varepsilon^{-1}\left[2\left(2 - a^2\right)\sin 2\phi - a^2 \sin 4\phi\right]. \tag{36}$$

By substituting (30) and (31) into (36) and equating the coefficients of similar powers of ε we obtain the following.

Order ε^{-1}:

$$\frac{\partial a_1}{\partial \overline{\phi}} + A_1 = 0.125\overline{a}\left(4 - \overline{a}^2\right) - 0.5\overline{a}\cos 2\overline{\phi} + 0.125\overline{a}^3 \cos 4\overline{\phi},$$
$$\frac{\partial \phi_1}{\partial \overline{\phi}} + B_1 = 0.25\left(2 - \overline{a}^2\right)\sin 2\overline{\phi} - 0.125\overline{a}^2 \sin 4\overline{\phi}. \tag{37}$$

Order ε^{-2}:

$$\frac{\partial a_2}{\partial \overline{\phi}} + A_2 = -\frac{\partial a_1}{\partial \overline{a}}A_1 - \frac{\partial a_1}{\partial \overline{a}}B_1 + 0.125a_1\left[4 - 3\overline{a}^2 - 4\cos 2\overline{\phi} + 3\overline{a}^2 \cos 4\overline{\phi}\right]$$
$$+ 0.5\overline{a}\phi\left[2\sin 2\overline{\phi} - \overline{a}^2 \sin 4\overline{\phi}\right],$$
$$\frac{\partial \phi_2}{\partial \overline{\phi}} + B_2 = -\frac{\partial \phi_1}{\partial \overline{a}}A_1 - \frac{\partial \phi_1}{\partial \overline{\phi}}B_1 - 0.25\overline{a}a_1\left(2\sin 2\overline{\phi} + \sin 4\overline{\phi}\right)$$
$$+ 0.5\phi_1\left[\left(2 - \overline{a}^2\right)\cos 2\overline{\phi} - \overline{a}^2 \cos 4\overline{\phi}\right]. \tag{38}$$

Equally, A_1 and B_1 by the long periodic terms on the right-hand side in (37) is obtained as

$$A_1 = 0.125\overline{a}\left(4 - \overline{a}^2\right), \quad B_1 = 0. \tag{39}$$

Hence, system (37) becomes

$$\frac{\partial a_1}{\partial \overline{\phi}} = -0.5\overline{a}\cos 2\overline{\phi} + 0.125\overline{a}^3 \cos 4\overline{\phi},$$
$$\frac{\partial \phi_1}{\partial \overline{\phi}} = 0.25\left(2 - \overline{a}^2\right)\sin 2\overline{\phi} - 0.125\overline{a}^2 \sin 4\overline{\phi}. \tag{40}$$

So, we obtain the following solutions

$$a_1 = -0.25\overline{a}\sin 2\overline{\phi} + 0.031\overline{a}^3 \sin 4\overline{\phi},$$
$$\phi_1 = -0.125\left(2 - \overline{a}^2\right)\cos 2\overline{\phi} + 0.031\overline{a}^2 \cos 4\overline{\phi}. \tag{41}$$

By using (39) and (41), the relation (38) becomes:

$$\frac{\partial a_2}{\partial \overline{\phi}} + A_2 = \text{short} - \text{period terms},$$
$$\frac{\partial \phi_2}{\partial \overline{\phi}} + B_2 = -0.125 + 0.188\overline{a}^2 - 0.043\overline{a}^4 \tag{42}$$
$$+ \text{short} - \text{period terms}.$$

Equally, A_2 and B_2 by the long periodic terms on the right-hand side in (42) is obtained as

$$A_2 = 0,$$
$$B_2 = -0.125 + 0.188\overline{a}^2 - 0.043\overline{a}^4. \tag{43}$$

Therefore, the solution up to the second approximation $(1/\varepsilon^2)$ becomes

$$u = a\cos\phi, \tag{44}$$

where

$$a = \overline{a} - 0.125\varepsilon^{-1}\overline{a}\left[\sin 2\overline{\phi} - 0.125\overline{a}^2 \sin 4\overline{\phi}\right] + o\left(\varepsilon^{-2}\right),$$

$$\phi = \overline{\phi} - 0.125\varepsilon^{-1}\left[\left(2 - \overline{a}^2\right)\cos 2\overline{\phi} - 0.25\overline{a}^2 \cos 4\overline{\phi}\right] + o\left(\varepsilon^{-2}\right),$$

$$\frac{\mathrm{d}\overline{a}}{\mathrm{d}t} = 0.125\varepsilon^{-1}\overline{a}\left(4 - \overline{a}^2\right) + o\left(\varepsilon^{-2}\right),$$

$$\frac{\mathrm{d}\overline{\phi}}{\mathrm{d}t} = 1 - 0.125\varepsilon^{-2}\left[1 - 1.5\overline{a}^2 + 0.344\overline{a}^4\right] + o\left(\varepsilon^{-3}\right). \tag{45}$$

3. Struble's Technique

Struble developed a technique for treating weak nonlinear oscillatory systems such as those governed by a second-order differential equation in terms of a small parameter [15]. We treat this equation in terms of the large parameter as follows:

$$\ddot{u} + \omega_0^2 u = \varepsilon^{-1} f(u, \dot{u}, t). \tag{46}$$

Let us consider the solution of (46) in terms of the large parameter ε, which takes the form

$$u = a\cos(\omega_0 t - \theta) + \sum_{n=1}^{N} \varepsilon^{-n} u_n(t) + o\left(\varepsilon^{-N-1}\right), \tag{47}$$

where a and θ are functions that are slowly changing with time.

As an example, we take Duffing's equation (19) and use (47) to give

$$\left[2a\omega_0\frac{\mathrm{d}\theta}{\mathrm{d}t} + \frac{d^2a}{\mathrm{d}t^2} - a\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2\right]\cos(\omega_0 t - \theta) + \left[-2\omega_0\frac{\mathrm{d}a}{\mathrm{d}t} + a\frac{d^2\theta}{\mathrm{d}t^2}\right.$$
$$\left. + 2\frac{\mathrm{d}a}{\mathrm{d}t}\frac{\mathrm{d}\theta}{\mathrm{d}t}\right]\sin(\omega_0 t - \theta) + \varepsilon^{-1}\left(\frac{d^2u_1}{\mathrm{d}t^2} + \omega_0^2 u_1\right) + \varepsilon^{-2}\left(\frac{d^2u_2}{\mathrm{d}t^2} + \omega_0^2 u_2\right) + \ldots$$
$$= -\varepsilon^{-1}a^3\cos^3(\omega_0 t - \theta) - 3\varepsilon^{-2}u_1 a^2\cos^2(\omega_0 t - \theta) + \ldots. \tag{48}$$

Considering the terms of order $O(\varepsilon^{-1})$ and equality of coefficients $\cos(\omega_0 t - \theta)$ and $\sin(\omega_0 t - \theta)$ on both sides, we get the so-called variation equations:

$$2a\omega_0\frac{\mathrm{d}\theta}{\mathrm{d}t} + \frac{d^2a}{\mathrm{d}t^2} - a\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2 = -0.75\varepsilon^{-1}a^3,$$
$$-2a\omega_0\frac{\mathrm{d}a}{\mathrm{d}t} + a\frac{d^2\theta}{\mathrm{d}t^2} + 2\frac{da}{\mathrm{d}t}\frac{d\theta}{\mathrm{d}t} = 0. \tag{49}$$

While the perturbed equations remain, so

$$\frac{d^2u_1}{\mathrm{d}t^2} + \omega_0^2 u_1 = -0.25a^3\cos 3(\omega_0 t - \theta). \tag{50}$$

For the first order of $(1/\varepsilon)$, equation (49) reduces to

$$\frac{\mathrm{d}a}{\mathrm{d}t} = 0,$$
$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -0.375\varepsilon^{-1}\left(\frac{a^2}{\omega_0}\right). \tag{51}$$

Hence,

$$a = a_0,$$
$$\theta = \left(-0.375\varepsilon^{-1}\left(\frac{a_0^2 t}{\omega_0}\right)\right) + \theta_0, \tag{52}$$

where a_o and θ_o are constants. If we consider θ and a are constants, the solution of (50) for the first order becomes

$$u_1 = 0.031\omega_0^{-2}a^3\cos 3(\omega_0 t - \theta). \tag{53}$$

Hence, the solution for the first-order $O(1/\varepsilon)$ becomes

$$u = a\cos(\omega_0 t - \theta) + 0.031\omega_0^{-2}\varepsilon^{-1}a^3\cos 3(\omega_0 t - \theta), \tag{54}$$

where a and θ are given from equation (52).

Satisfying the solution for second order, we use (53) for calculating the term:

$$-3\varepsilon^{-2}u_1 a^2\cos^2(\omega_0 - \theta) = -0.023\omega_0^{-2}\varepsilon^{-2}a^5\left[\cos(\omega_0 t - \theta)\right.$$
$$\left. + 2\cos 3(\omega_0 t - \theta) + \cos 5(\omega_0 t - \theta)\right]. \tag{55}$$

We note that the expression $((d^2u_1/\mathrm{d}t^2) + \omega_0^2 u_1)$ contains a term of order $O(1/\varepsilon)$ in the following form:

$$0.563\omega_0^{-2}a^3\frac{\mathrm{d}\theta}{\mathrm{d}t}\cos 3(\omega_0 t - \theta). \tag{56}$$

Now, we consider the order $O(1/\varepsilon^2)$ leads us to the variation equations:

$$2a\omega_0\frac{\mathrm{d}\theta}{\mathrm{d}t} + \frac{d^2a}{\mathrm{d}t^2} - a\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2 = -0.75\varepsilon^{-1}a^3 - 0.023\omega_0^{-2}\varepsilon^{-2}a^5$$
$$-2a\omega_0\frac{\mathrm{d}a}{\mathrm{d}t} + a\frac{d^2\theta}{\mathrm{d}t^2} + 2\frac{\mathrm{d}a}{\mathrm{d}t}\frac{\mathrm{d}\theta}{\mathrm{d}t} = 0, \tag{57}$$

and the perturbed equation is

$$\frac{d^2u_2}{dt^2} + \omega_0^2 u_2 = -0.023\omega_0^{-2}a^5\left[2\cos 3\left(\omega_0 t - \theta\right) + \cos 5\left(\omega_0 t - \theta\right)\right] - 0.563\omega_0{}^{-1}\varepsilon^{-1}a^3\frac{d\theta}{dt}\cos 3\left(\omega_0 t - \theta\right). \tag{58}$$

The solution of equation (57) can be obtained from the recurrence principle of (52) to obtain

$$a = a_0,$$
$$\theta = -0.375\omega_0^{-1}\varepsilon^{-1}a_0^2 t + 0.059\omega_0^{-3}\varepsilon^{-2}a_0^4 t + \theta_0 + o\left(\varepsilon^{-3}\right), \tag{59}$$

where a_o and θ_o are constants. Substituting by $(d\theta/dt)$ from (52) into (58) and solving the resulted equation, we obtain

$$u_2 = -0.001\omega_0^{-4}a^5\left[21\cos 3\left(\omega_0 t - \theta\right) - \cos 5\left(\omega_0 t - \theta\right)\right]. \tag{60}$$

Thus, we obtain the solution up to the second approximation in the form

$$u = a\cos\left(\omega t - \theta_0\right) + 0.031\varepsilon^{-1}a^3\omega_0^{-2}\left(1 - 0.656\varepsilon^{-1}\omega_0^{-2}a^2\right)\cos 3\left(\omega t - \theta_0\right) + 0.001\varepsilon^{-2}a^5\omega_0^{-4}\cos 5\left(\omega t - \theta_0\right) + o\left(\varepsilon^{-3}\right), \tag{61}$$

where

$$\omega = \omega_0\left(1 + 0.375\varepsilon^{-1}a^2\omega_0^{-2} - 0.059\varepsilon^{-2}a^4\omega_0^{-4}\right) + o\left(\varepsilon^{-3}\right). \tag{62}$$

The periodic time $T = (2\pi/\omega)$ is given in the form

$$T = 2\pi\omega_0^{-1} + \alpha\left(\varepsilon^{-1}\right), \tag{63}$$

where the correction of the period is obtained in the form

$$\alpha\left(\varepsilon^{-1}\right) = 2\pi\omega_0^{-1}\left[-0.375\varepsilon^{-1}a^2\omega_0^{-2} + 0.199\varepsilon^{-2}a^4\omega_0^{-4} + o\left(\varepsilon^{-3}\right)\right]. \tag{64}$$

4. The Krylov–Bogoliubov–Mitropolski Technique

This technique is devoted to finding approximated periodic solutions using the small parameter as in [7–11]. In our work, we modify the approximated expansions using the large parameter to become

$$u = a\cos\psi + \sum_{n=1}^{N}\varepsilon^{-n}u_n(a,\psi) + o\left(\varepsilon^{-N-1}\right), \tag{65}$$

where each u_n is a periodic function in ψ, and its period 2π, a, and ψ are imposed to change with time as follows:

$$\frac{da}{dt} = \sum_{n=1}^{N}\varepsilon^{-n}A_n(a) + o\left(\varepsilon^{-N-1}\right), \tag{66}$$

$$\frac{d\psi}{dt} = \omega_o + \sum_{n=1}^{N}\varepsilon^{-n}\psi_n(a) + o\left(\varepsilon^{-N-1}\right), \tag{67}$$

where the functions u_n, A_n, and ψ_n are arbitrary so that equations (65) to (67) satisfy the differential equation (9). The following derivatives series are obtained:

$$\frac{d}{dt} = \frac{da}{dt}\frac{\partial}{\partial a} + \frac{d\psi}{dt}\frac{\partial}{\partial\psi},$$
$$\frac{d^2}{dt^2} = \left(\frac{da}{dt}\right)^2\frac{\partial^2}{\partial a^2} + \frac{d^2a}{dt^2}\frac{\partial}{\partial a} + 2\frac{da}{dt}\frac{d\psi}{dt}\frac{\partial^2}{\partial a\,\partial\psi} + \left(\frac{d\psi}{dt}\right)^2\frac{\partial^2}{\partial\psi^2} + \frac{d^2\psi}{dt^2}\frac{\partial}{\partial\psi}; \tag{68}$$

$$\frac{d^2a}{dt^2} = \frac{da}{dt}\sum_{n=1}^{N}\varepsilon^{-n}\frac{dA_n}{da} = \varepsilon^{-2}A_1\frac{dA_1}{da} + o\left(\varepsilon^{-3}\right),$$
$$\frac{d^2\psi}{dt^2} = \frac{da}{dt}\sum_{n=1}^{N}\varepsilon^{-n}\frac{d\psi_n}{da} = \varepsilon^{-2}A_1\frac{d\psi_1}{da} + o\left(\varepsilon^{-3}\right). \tag{69}$$

We will visualize this method in the following example.

4.1. Duffing Equation. We consider a nonlinear oscillator (19), using equations (65) to (69), and equate the coefficients of similar powers to $(1/\varepsilon)$, we obtain

$$\omega_o\left(\frac{\partial^2 u_1}{\partial\psi^2}+u_1\right)=2\left(\psi_1 a\cos\psi+A_1\sin\psi\right)-a^3\omega_o^{-1}\cos^3\psi, \tag{70}$$

$$\begin{aligned}\omega_o^2\left(\frac{\partial^2 u_2}{\partial\psi^2}+u_2\right)=&\left[a\left(2\omega_o\psi_2+\psi_1^2\right)-A_1\frac{\mathrm{d}A_1}{\mathrm{d}a}\right]\cos\psi+\left[2\left(\omega_o A_2+A_1\psi_1\right)\right.\\&\left.+aA_1\frac{\mathrm{d}\psi_1}{\mathrm{d}a}\right]\sin\psi-3u_1a^2\cos^2\psi-2\omega_o\left(\psi_1\frac{\partial^2 u_1}{\partial\psi^2}+A_1\frac{\partial^2 u_1}{\partial a\,\partial\psi}\right).\end{aligned} \tag{71}$$

Let u_1 be periodic, and the singular terms in the right-hand side of (70) must be deleted. Since

$$\cos^3\psi=0.75\left(3\cos\psi+\cos 3\psi\right), \tag{72}$$

then equation (70) gives

$$\begin{aligned}A_1&=0,\\\psi_1&=0.375a^2\omega_o^{-1}.\end{aligned} \tag{73}$$

The solution of equation (70) becomes

$$u_1=0.031a^3\omega_o^{-2}\cos 3\psi. \tag{74}$$

Substituting (74) into (71) gives

$$\begin{aligned}\omega_o^2\left(\frac{\partial^2 u_2}{\partial\psi^2}+u_2\right)=&\left(2\omega_o\psi_2+0.117a^4\omega_o^{-2}\right)a\cos\psi+2\omega_o A_2\sin\psi\\&+0.008a^5\omega_o^{-2}\left(21\cos 3\psi-3\cos 5\psi\right).\end{aligned} \tag{75}$$

Canceling the singular terms yields

$$\begin{aligned}A_2&=0,\\\psi_2&=-0.059a^4\omega_o^{-3}.\end{aligned} \tag{76}$$

The solution of (75) becomes

$$u_2=-0.001a^5\omega_o^{-4}\left(21\cos 3\psi-\cos 5\psi\right). \tag{77}$$

Therefore, the second approximated solution u is given in the following form:

$$u=a\cos\psi+0.031\varepsilon^{-1}a^3\omega_0^{-2}\cos 3\psi-0.001\varepsilon^{-2}a^5\omega_0^{-4}\left(21\cos 3\psi-\cos 5\psi\right)+o\left(\varepsilon^{-3}\right), \tag{78}$$

where

$$\begin{aligned}&\frac{\mathrm{d}a}{\mathrm{d}t}=0\text{ or }a=a_0=\text{const.},\\&\frac{\mathrm{d}\psi}{\mathrm{d}t}=\omega_0+0.375\varepsilon^{-1}a^2\omega_0^{-1}-0.059\varepsilon^{-2}a^4\omega_0^{-3}+o\left(\varepsilon^{-3}\right),\\&\psi=\omega_0\left[1+0.375\varepsilon^{-1}a^2\omega_0^{-2}-0.059\varepsilon^{-2}a^4\omega_0^{-4}\right]t+\psi_0+o\left(\varepsilon^{-3}\right),\end{aligned} \tag{79}$$

where ψ_o is constant.

5. Multiple Scale Method

This technique is considered for solving the differential equations applying the small parameter [1, 16]. In our work, we modify this method to be suitable for the large parameter technique.

From equation (46), we put $\omega_o=1$, and we obtain

$$\ddot{u} + u = \varepsilon^{-1} f(u, \dot{u}, t). \tag{80}$$

To have accurate and regular expansion over all periods for the solution of this equation, we should be finding an approximated representation of the function $u(t, \varepsilon^{-1})$ as shown in the form

$$u\left(t, \varepsilon^{-1}\right) = \sum_{n=0}^{N} \varepsilon^{-n} u_n^*(t_0, \ldots, t_N) + o\left(\varepsilon^{-N-1}\right), \tag{81}$$

where

$$t_m = \varepsilon^{-m} t, \quad (m = 0, 1, 2, \ldots, N). \tag{82}$$

Therefore, the total differential of (80) gives

$$\frac{d}{\mathrm{d}t} = \sum_{n=o}^{N} \varepsilon^{-n} \frac{\partial}{\partial t_n}. \tag{83}$$

Substituting expansions (81) and (83) into equation (80), then equalizing the coefficients of equal powers for ε^{-1}on both sides, and by solving the resulting equations, the solutions are obtained. Such solutions contain arbitrary functions in the variables $t_1, \ldots, t_N$ and will be determined with a condition u_o^* and (u_n^*/u_{n-1}^*) which are bounded to each t_n.

Now, we will apply the modified method to the following example.

5.1. Duffing Equation. Consider the Duffing equation in terms of the large parameter which takes the form

$$\frac{d^2 x}{\mathrm{d}t^2} + \omega_0^2 x + \varepsilon^{-1} x^3 = 0. \tag{84}$$

Taking

$$x = \sum_{n=o}^{2} \varepsilon^{-n} x_n(t_o, t_1, t_2) + o\left(\varepsilon^{-3}\right), \tag{85}$$

substituting (85) in (84), and equating the like powers of ε^{-1} to zero, we obtain

$$D_0^2 x_0 + \omega_0^2 x_0 = 0, \tag{86}$$

$$D_0^2 x_1 + \omega_0^2 x_1 + 2D_0 D_1 x_0 + x_0^3 = 0, \tag{87}$$

$$D_0^2 x_2 + \omega_0^2 x_2 + 2D_0 D_1 x_1 + \left(2D_0 D_2 + D_1^2\right) x_0 + 3x_0^2 x_1 = 0, \tag{88}$$

where $D_n = (\partial/\partial t_n)$.

The solution of equation (86) is

$$x_0 = A(t_1, t_2) e^{i\omega_0 t} + \overline{A}(t_1, t_2) e^{-i\omega_0 t}, \tag{89}$$

where $\overline{A}$ is the complex conjugate of A. Then, equation (87) becomes as follows:

$$\begin{aligned} D_0^2 x_1 + \omega_0^2 x_1 = {} & \left[2i\omega_0 D_1 \overline{A} - 3A\overline{A}^2\right] e^{-i\omega_0 t} - \left[2i\omega_0 D_1 A + 3A^2 \overline{A}\right] e^{i\omega_0 t} \\ & - A^3 e^{3i\omega_0 t} - \overline{A}^3 e^{-3i\omega_0 t}. \end{aligned} \tag{90}$$

Since (x_1/x_0) is finite for each t, the singular terms must be neglected. So,

$$\begin{aligned} 2i\omega_0 D_1 A + 3A^2 \overline{A} &= 0, \\ 2i\omega_0 D_1 \overline{A} - 3A\overline{A}^2 &= 0, \end{aligned} \tag{91}$$

thus

$$\begin{aligned} A &= A_1(t_2) e^{\left(3iA_1 \overline{A}_1 t_1 / 2\omega_0\right)}, \\ x_1 &= B e^{i\omega_0 t} + \overline{B} e^{-i\omega_0 t} + 0.125\omega_0^{-2}\left[A^3 e^{3i\omega_0 t} + \overline{A}^3 e^{-3i\omega_0 t}\right]. \end{aligned} \tag{92}$$

Using x_0 and x_1 into (88) yields

$$D_0^2 x_2 + \omega_0^2 x_2 = -0.375\omega_0^{-2}\left[A^5 e^{5i\omega_0 t} + \overline{A}^5 e^{-5i\omega_0 t}\right] + 3A^2\left[0.875\omega_0^{-2}A^2\overline{A} - B\right]e^{3i\omega_0 t} + 3\overline{A}^2\left[0.875\omega_0^{-2}A\overline{A}^2 - \overline{B}\right]e^{-3i\omega_0 t} + Qe^{i\omega_0 t} + \overline{Q}e^{-i\omega_0 t}, \tag{93}$$

where

$$Q = -2i\omega_0\left(D_1 B + D_2 A\right) - 3A\left(A\overline{B} + 2B\overline{A}\right) + 1.875\omega_0^{-2}A^3\overline{A}^2. \tag{94}$$

By eliminating the singular terms, therefore (x_2/x_1) will be finite for all t; then,

$$B = 0, \qquad 2i\omega_0 D_2 A_1 = 1.875\omega_0^{-2}A_1^3\overline{A}_1^2. \tag{95}$$

Therefore,

$$A_1 = A_2 e^{-0.938iA_2^2\overline{A}_2^2\omega_0^{-3}t_2}. \tag{96}$$

A solution x_2 regardless of the homogeneous solution is given:

$$x_2 = 0.016\omega_0^{-4}\left[A^5 e^{5i\omega_0 t} + \overline{A}^5 e^{-5i\omega_0 t} - 21A\overline{A}\left(A^3 e^{3i\omega_0 t} + \overline{A}^3 e^{-3i\omega_0 t}\right)\right]. \tag{97}$$

Finally, we get the solution x in terms of the large parameter up to the second approximation of $(1/\varepsilon)$ in the form

$$x = a\cos(\omega t + \phi) + 0.031\varepsilon^{-1}a^3\omega_0^{-2}\left(1 - 0.656\varepsilon^{-1}a^2\omega_0^{-2}\right)\cos 3(\omega t + \phi) + 0.001\varepsilon^{-2}a^5\omega_0^{-4}\cos 5(\omega t + \phi), \tag{98}$$

where $A = 0.5ae^{i\phi}$ and $\omega = \omega_0 + 0.375\varepsilon^{-1}a^2\omega_0^{-1} - 0.059\varepsilon^{-2}a^4\omega_0^{-3}$. The obtained solution has the same period (63).

6. Conclusion

In this paper, we modified some of the perturbation methods that are considered ones of the innovations of mathematicians and physicists for finding approximated solutions to important problems. We cannot find a complete solution to these problems. Nayfeh, Poincare, and Krylov are among the prominent scientists in this field as they dealt with many problems in mathematics and physics which are applied in many branches, such as classical mechanics, flexible body mechanics, and many other branches. All these methods depend on a small parameter that was defined for each problem.

The study of such methods shows that the large parameter method is the easiest, most accurate, and one of the methods most used to solve many important problems efficiently. The large parameter method has an advantage over the other methods because it solves the problem in a new domain when it fails all other methods for solving the problem in such a domain. One of the most important application is when we study the slow spin motion of a rigid body in a Newtonian field of force under an external moment [4], the rotational motion of a heavy solid in a uniform gravity field [3], or the gyroscopic motions with a sufficiently small angular velocity component about the major or the minor axis of the ellipsoid of inertia. There are many applications of this technique in aerospace science, satellites, navigations, antennas, and solar collectors. This technique is also useful in all perturbed problems in physics and mechanics, for example, the perturbed pendulum motions and the perturbed mechanical systems.

Data Availability

No data were used to support the findings of the study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. H. Nayfeh, *Introduction to Perturbation Techniques*, pp. 360–364, WILEY-VCH Verlag GmbH & Co. KGaA, Hoboken, NJ, USA, 2011.

[2] I. A. Arkhangel'skii, *Construction of Periodic Solutions for the Euler-Poisson Equations by Means of Power Series Expansions Containing a Small Parameter*, Colloquia Mathematica Societatis Janos Bolyai, Keszthely, Hungary, 1975.

[3] A. I. Ismail, "Solving a problem of rotary motion for a heavy solid using the large parameter method," *Advances in Astronomy*, vol. 2020, Article ID 2764867, , 2020.

[4] A. I. Ismail, "The slow spinning motion of a rigid body in Newtonian field and external torque," *Advances in Astronomy*, vol. 2020, Article ID 4179590, , 2020.

[5] A. P. Proskuriakov, "Periodic oscillations of a quasilinear autonomous system with two degrees of freedom," *Public Money & Management*, vol. 24, no. 6, pp. 1103–1109, 1960.

[6] Z. H. A. Sartabanov and B. Z. H. Omarova, *On Multi-Periodic Solutions of Quasilinear Autonomous Systems with an Operator of Differentiation on the Lyapunov's Vector Field*, K. Zhubanov Aktobe Regional State University, Aktobe, Kazakhstan, 2019.

[7] N. N. Bogoliubov and Y. A. Mitropolski, *Asymptotic Methods in the Theory of Non-linear Oscillations*, Gordon and Breach, New York, NY, USA, 1961.

[8] A. I. Ismail, "On the application of Krylov-bogoliubov-mitropolski technique for treating the motion about a fixed point of a fast spinning heavy solid," *Zeitschrift für Wirtschaftsgeographie*, vol. 20, pp. 205–208, 1996.

[9] A. I. Ismail, "Periodic solutions of equations of motion of a heavy solid applying Krylov-Bogoliubov-Mitropolski method," *Journal of Computational and Applied Mathematics*, vol. 114, no. 2, pp. 345–359, 2000.

[10] T. S. Amer, A. I. Ismail, and W. S. Amer, "Application of the krylov-bogoliubov-mitropolski technique for a rotating heavy solid under the influence of a gyrostatic moment," *Journal of Aerospace Engineering*, vol. 25, no. 3, pp. 421–430, 2012.

[11] T. S. Amer and I. M. Abady, "On the application of KBM method for the 3-D motion of asymmetric rigid body," *Nonlinear Dynamics*, vol. 89, no. 3, pp. 1591–1609, 2017.

[12] Y. A. Mitropolski, "Averaging methods in non-linear mechanics," *International Journal of Non-Linear Mechanics*, vol. 2, no. 1, pp. 69–96, 1967.

[13] A. I. Ismail, T. S. Amer, S. A. El Banna, and M. A. El-Ameen, "Electromagnetic gyroscopic motion," *Journal of Applied Mathematics*, vol. 2012, Article ID 12645, , 2012.

[14] T. S. Amer, "The rotational motion of the electromagnetic symmetric rigid body," *Applied Mathematics & Information Sciences*, vol. 10, no. 4, pp. 1453–1464, 2016.

[15] F. Lakrad and M. M. Charafi, "Perturbation methods and the Melnikov functions for slowly varying oscillators," *Chaos, Solitons & Fractals*, vol. 25, no. 3, pp. 675–680, 2005.

[16] P. K. Jakobsen, "Introduction to the method of multiple scales," 2016, https://arxiv.org/abs/1312.3651.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 6686992, 8 pages
https://doi.org/10.1155/2021/6686992

Research Article

Supply Interruption Supply Chain Network Model with Uncertain Demand: An Application of Chance-Constrained Programming with Fuzzy Parameters

Haidong Guo,[1,2] **Shengyu Wang**,[3] **and Yu Zhang**[4]

[1]*School of Economics, Zhejiang University of Technology, Hangzhou 310023, China*
[2]*College of Education, Zhejiang University of Technology, Hangzhou 310023, China*
[3]*School of Finance and Trade, Wenzhou Business College, Wenzhou 325035, Zhejiang, China*
[4]*Department of Media and Communications, London School of Economics and Political Science, London, UK*

Correspondence should be addressed to Haidong Guo; haidongguo1976@163.com

Received 19 November 2020; Revised 27 December 2020; Accepted 17 January 2021; Published 8 February 2021

Academic Editor: Stefania Tomasiello

The downstream supply interruption of manufacturers is a disaster for the company when the demand is uncertain in the market; a fuzzy programming with fuzzy parameters model of supply interruption supply chain network is established by simulating market operation rules. The aim of the current study is to build a fuzzy chance-constrained programming method which is developed for supporting the uncertainty of demand. This method ensured that the fuzzy constraints can be satisfied at specified confidence levels, leading to cost-effective solutions under acceptable risk magnitudes. Finally, through the case of the electronic product manufacturing enterprise, the feasibility and effectiveness of the proposed model are verified by adopting a sensitivity analysis of capacity loss level and minimizing objective function. Numerical simulation shows that selecting two manufacturing centers can effectively reduce the supply chain cost and maintain business continuity.

1. Introduction

In the actual production process, the uncertainty of demand caused by the individualization and diversity requirements of consumers often impacts the productivity of manufacturers [1–3]. According to the Japan Economic Daily, Foxconn's original 60 iPhone XR production lines, in November 2018, due to insights into changes in product demand in the market, 45 production lines were decided to start and the remaining 15 production lines were suspended indefinitely because the manufacturer's production line usually changed according to user demand. This crisis is manifested in the risk of disruption to the supply capacity of the end customer. For example, in 2019, Xiao Mi's new mobile phone Xiaomi 9 was sold out in a minute after the release of Jing Dong due to insufficient stocking, which seriously affected the profit of Xiaomi enterprises. The Vivo IQOO mobile phone is also popular, but the supply is sufficient, effectively avoiding the risk of supply interruption. Analysis of the underlying causes shows that the reliability of the manufacturing centers of the two companies is the main source of the supply disruption crisis. Among them, Xiaomi Company is organized by Foxconn and Inventec Appliances Corp. Foundry companies have their own interests, usually not only for Xiaomi's enterprises, but also preferentially produce orders with large profits, so there will be a risk of disruption of Xiaomi's mobile phone supply. The Vivo Company has four production bases in Dongguan, Chongqing, New Delhi, India, and Jakarta, Indonesia. 122 modern production lines can produce 10 million mobile phones per month, and the manufacturing center of Vivo is dedicated to the production of Vivo mobile phones and service, thus providing a stable source of supply for mobile phone sales.

The solution to supply disruption is usually to form different supply chain risk management models based on the

analysis of the company's competitive strategy [4–6]. Because the supply chain flexible design can effectively solve the integration of market and enterprise, it is of great significance to deal with potential supply chain risks [7]. Often, the company's flexible design of the supply chain after determining its strategic target market requires attention to two main characteristics: supply chain scope (local or global) and competitive priority (responsiveness or cost reduction) [8–10]. At this stage, it is important to understand the supply chain vulnerability caused by supply disruptions, which will help to achieve a dynamic adaptive elastic supply chain. Figure 1 illustrates the dynamic process of supply chain resiliency design in this paper and demonstrates the optimization process of supply chain risk management to alleviate supply chain risk. This process is mainly achieved through supply chain resiliency design [11–14].

At present, the research on the risk of supply interruption is rich, mainly including multisupplier strategy research [15–17]; the establishment of emergency inventory [18, 19]; alternative suppliers [20, 21]; contract theory to reduce futures commodity risks [22]; [23–25]. Facility safety is improved to cope with the supply disruption crisis caused by natural risks [26, 27], but the above research usually lacks a flexible mechanism and cannot respond to supply disruption crisis according to supply chain structure adjustment and different situational flexibility change strategies [28–31]. However, research methods for uncertainty requirements include the use of fuzzy programming or robust optimization to deal with uncertainty, but these methods usually require a large number of samples, but in fact, the number of samples is limited and only can get the range of sample values rather than the exact value [8, 32, 33].

Therefore, this paper constructs a three-tier supply chain network model, consisting of one supplier, multiple manufacturers, multiple customers, and a manufacturing company and studies the risk of supply disruption when demand is uncertain [34–40]. The fuzzy number of uncertain demand is converted into the determined risk of supply interruption risk, and the logistics supply is used to coordinate the supply capacity of reliable and unreliable manufacturing centers for supply chain flexibility design, to meet the needs of users while maintaining the stability of the supply chain and reducing the risk of supply [41–45].

2. Model Assumptions and Construction

2.1. Description of the Problem. A three-tier supply chain network of suppliers, multiple manufacturers, and multiple-end customers, is built as the demand is uncertain, as shown in Figure 2. Assume that the parts in the market are supplied by one supplier, and the manufacturing center organizes the production according to the customer's needs. The reliable manufacturing center produces according to its production capacity to meet the needs of customers. The unreliable manufacturer can adjust the production plan according to its interests and may not meet the needs of the customer. This means that the quantity of products provided by reliable manufacturing centers is stable, while the quantity of products provided by unreliable manufacturers is unstable. The quantity of products required, and the shortage can be met by the reliable manufacturing center through the coordination of logistics functions from the reliable manufacturing center.

Figure 2 shows two manufacturing centers, which are two types of manufacturing centers. In a supply chain, each type can have multiple manufacturing centers. Among them, the manufacturing center cannot exceed the maximum production capacity of the customer, and the supplier's supply capacity and the production capacity of the manufacturing center are all certain values; the coordination of transportation products between the manufacturing centers does not affect the customer's needs. The location and number of the center are known; each customer only picks up one center for delivery, the reliable manufacturing center does not interrupt, the unreliable manufacturing center will interrupt with a certain probability, and some of the supply capacity will be lost when the interruption occurs, but which can be complemented by a reliable manufacturing center.

2.2. Symbols and Meanings

2.2.1. Collection Definition

I: reliable production center number, $i \in I$

J: unreliable manufacturing center number, $j \in J$

L: all manufacturing center numbers, $l \in L$ 且 $L \in (I \cup J)$

K: customer point number, $k \in K$

2.2.2. Parameter Definition

$\otimes d_k$: fuzzy demand for the kth customer

d_{ab}: transportation cost from manufacturing center a to manufacturing center b or customer

d_{pi}: unit transportation cost from supplier to reliable manufacturing center

d_{pj}: unit transportation cost from supplier to unreliable manufacturing center

γ_1: manufacturing center 1 distribution capacity

CR_i: start-up cost of a reliable manufacturing center

CR_j: the cost of starting an unreliable manufacturing center

qj: unreliable distribution center j

pj: loss ratio of capacity after unreliable distribution center j interruption

2.2.3. Decision Variables

U_i: supplier flow to manufacturing center i

U_j: supplier flow to manufacturing center j

FR_{ik}: reliable manufacturing center to customer k

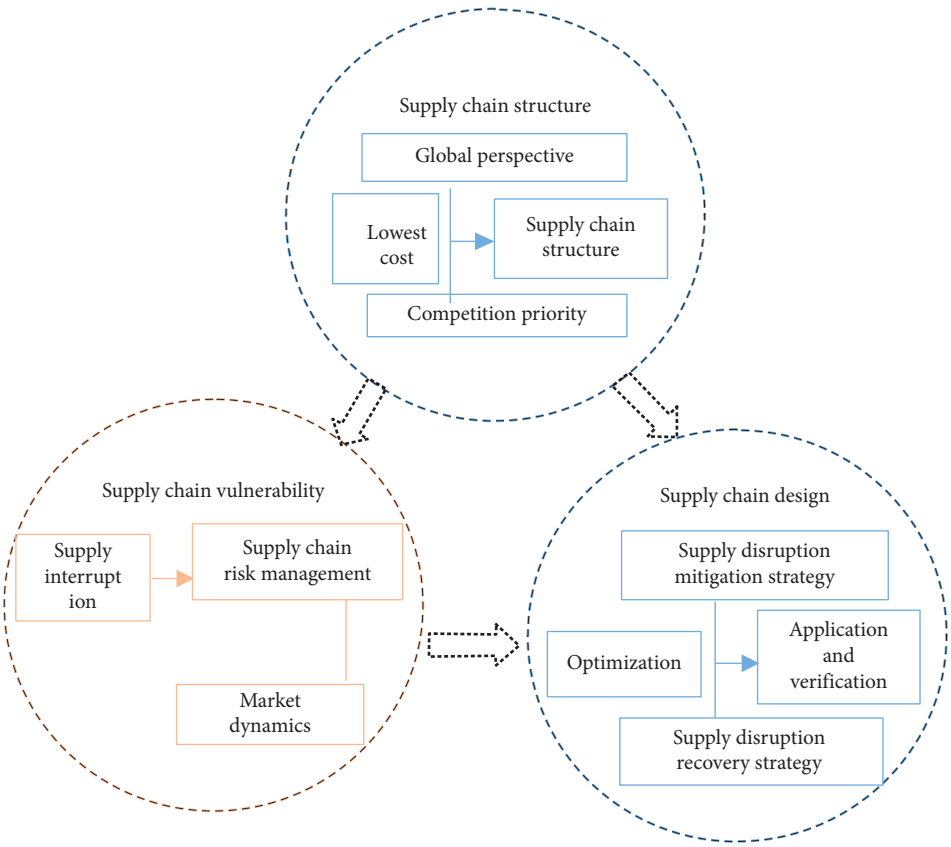


Figure 1: Dynamically built supply chain flexibility concept map.

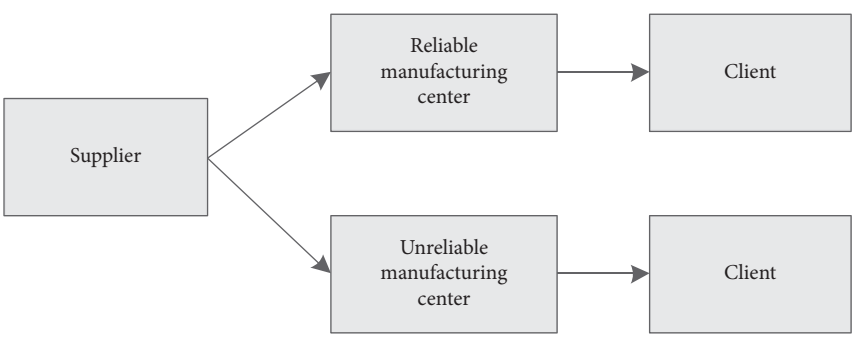


Figure 2: Three-level global supply chain network structure.

FU_{jk}: unreliable manufacturing center to customer k

T_{ij}: the number of deliveries from the reliable manufacturing center i to the unreliable manufacturing center j after the interruption occurred

YR_i: binary variable, 1 means reliable manufacturing center open; otherwise, 0

YU_j: binary variable, 1 means unreliable manufacturing center open; otherwise, 0

AR_{ik}: binary variable, 1 represents the customer assigned to a reliable manufacturing center; otherwise, 0

AU_{jk}: binary variable, 1 represents the customer assigned to the unreliable manufacturing center; otherwise, 0

2.3. Mathematical Modelling

$$\min \sum_i CR_i YR_i + \sum_j CU_j YU_j + \sum_i U_i d_{pi} + \sum_j U_j d_{pj} + \sum_i \sum_k \otimes d_k AR_{ik} d_{ik} + \sum_j \sum_k \otimes d_k AU_{jk} d_{jk} + \sum_i \sum_j q_j T_{ij} d_{ij}, \tag{1}$$

$$s.t. \sum_i AR_{ik} + \sum_j AU_{jk} = 1, \quad \forall k \in K, \tag{2}$$

$$YR_l + YU_l \leq 1, \quad \forall l \in L, \tag{3}$$

$$\sum_i YR_i \geq 1, \tag{4}$$

$$AR_{ik} \leq YR_i, \tag{5}$$

$$AU_{jk} \leq YU_j, \tag{6}$$

$$\sum_j FU_{jk} \bullet AU_{jk} \leq T_{ij} + \left(1 - p_j\right) \bullet \gamma_j YU_j, \tag{7}$$

$$T_{ij} + \left(1 - p_j\right) \bullet \gamma_j YU_j \leq \gamma_j YU_j, \tag{8}$$

$$\sum_k FU_{jk} \bullet AU_{jk} \geq \sum_k \otimes d_k \bullet AU_{jk}, \tag{9}$$

$$\sum_k FU_{ik} \bullet AU_{ik} \geq \sum_k \otimes d_k \bullet AU_{ik}, \tag{10}$$

$$YR_i, YU_j, AR_{ik}, AU_{jk} \in \{0, 1\}. \tag{11}$$

Among them, formula (1) is an objective function, which indicates the start-up cost of the manufacturing center, the transportation cost between the facilities, and the transshipment cost from the reliable manufacturing center to the unreliable manufacturing center, with the goal of minimizing the total cost. The constraint condition (2) indicates that a customer can only be assigned to one manufacturing center to meet the demand; formula (3) indicates that only one reliable manufacturing center or unreliable manufacturing center can be constructed at one location; and (4) indicates at least one reliable manufacturing center; formulas (5) and (6) indicates that the customer is assigned to a reliable manufacturing center or an unreliable manufacturing center; formulas (7) and (8) indicate that the flow rate of the material flowing out in the unreliable distribution center is not greater than the sum of the inflow amount and the amount of the inflow. And, less than its production capacity, equations (9) and (10) indicate that customer requirements should all be satisfied; equation (11) represents the value range constraint of the control variables.

2.4. Conversion of Fuzzy Constraints. The objective function in the above model contains uncertain parameters, and it is difficult to solve. Therefore, according to the random chance-constrained programming method in the uncertainty theory proposed by Liu and its followers [46, 47], the fuzzy number model with the customer's uncertain demand is transformed into a deterministic model. The fuzzy demand for the customer $\otimes d_k$ is expressed as a triangular fuzzy number (d_{j1}, d_{j2}, d_{j3}). The customer demand function $\mu_{\otimes d_k}(x)$ can be expressed as

$$\mu_{\otimes d_k}(x) = \begin{cases} \dfrac{x - d_{j1}}{d_{j2} - d_{j1}}, & \text{when } d_{j1} \leq x \leq d_{j2}, \\ \dfrac{x - d_{j3}}{d_{j2} - d_{j3}}, & \text{when } d_{j2} \leq x \leq d_{j3}, \\ 0, & \text{otherwise.} \end{cases} \tag{12}$$

For a given confidence level θ $(0 \leq \theta \leq 1)$, according to the equivalence theorem in the stochastic chance-constrained programming, d_k can be obtained if and only if the following conditions are met:

$$\begin{cases} d_k \geq (1 - \theta) d_{j1} + \theta d_{j2}, \\ d_k \leq (1 - \theta) d_{j2} + \theta d_{j3}. \end{cases} \tag{13}$$

3. Case Analysis

For an electronic product manufacturing enterprise, due to business development needs, the manufacturing center is deployed in the world for the production of electronic products. The manufacturing center is divided into two categories. One is a reliable manufacturing center, which is invested by itself. The production capacity of the manufacturing center will not be interrupted because of the uncertainty of demand and the output produced by it is stable. The other type is entrusted to other enterprises to process. This category is an unreliable manufacturing center. The production capacity of the manufacturing center will be adjusted due to the interests of the processing enterprise itself. Suppose that there is one supplier, four manufacturing centers (including reliable manufacturing centers and unreliable manufacturing centers), and six end customers. The confidence level of customer demand uncertainty is $\theta = 0.8$, assuming an outage probability of 0.15, and the capacity loss ratio of an unreliable manufacturing center is 0.3. The other parameters are listed in Table 1.

Substitute the parameters into the model and solve them with MATLAB software. The results are shown in Table 2.

As can be seen from Table 2, the objective function is 115,586 yuan, of which the manufacturing center with the opening numbers 1, 3, and 4 is a reliable manufacturing center invested and built by electronic product manufacturing enterprises, and the manufacturing center numbered 2 is entrusted to other enterprises for processing by electronic product manufacturing enterprises, which is an unreliable manufacturing center. The total cost of starting a reliable manufacturing center and an unreliable manufacturing center is 102,139 yuan, the transportation cost is 13,397 yuan, and the transportation cost is 65 yuan. The transfer volume of the reliable manufacturing center 3 to

Table 1: Partial parameter range table.

Parameter	Range	Parameter	Range
γ_i	Uniform (300, 520)	CU_j	Uniform (18000, 26000)
θ	0.8	CR_i	Uniform (21600, 31200)
d_{ab}	(1, 6, 11)	—	—

Table 2: Results of the study.

Objective function value	Manufacturing center cost	Transportation cost	Transshipment cost	YR_i	YU_i
115,585	102,139	13,381	65	1, 3, 4	2

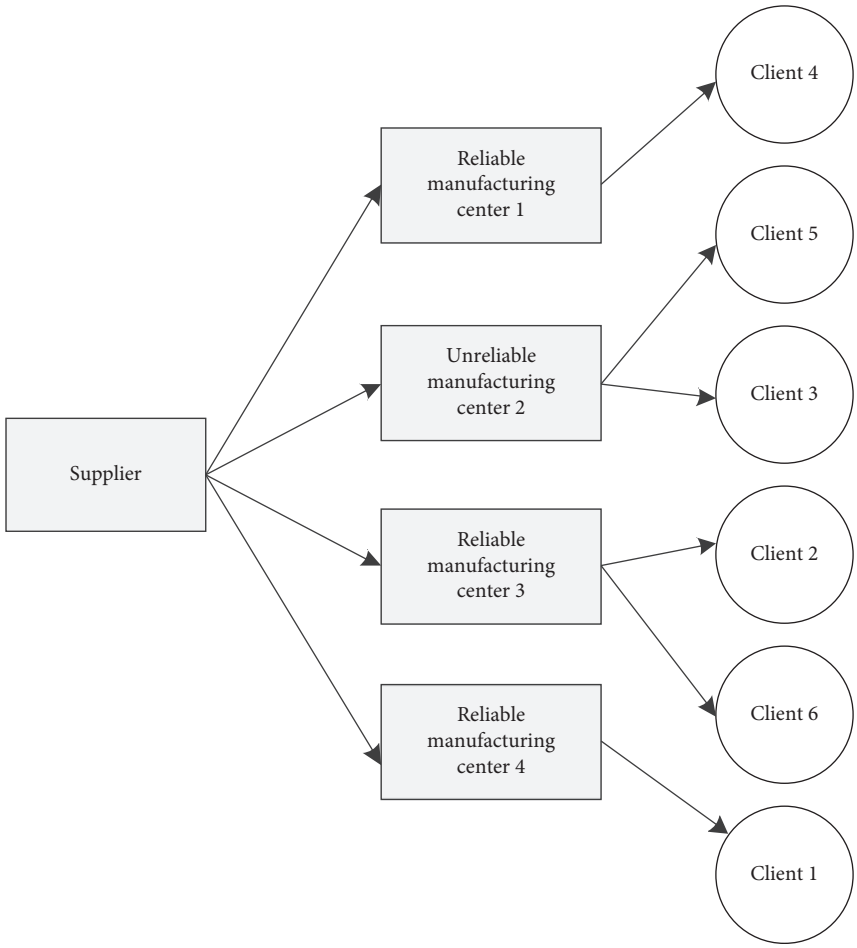


Figure 3: Intersite logistics transfer map.

the unreliable manufacturing center 2 is 56.3, as shown in Figure 3.

Electronic manufacturing companies choose to invest in building manufacturing centers or entrust other foundry companies to make decisions between processing products. When entrusting other processing enterprises to OEM, they will face the interests of other foundry companies, there is a risk of supply interruption, resulting in loss of processing capacity, and the products cannot be delivered on schedule. In order to better analyze the relationship between capacity loss parameters and product cost expenditure, sensitivity analysis of capacity loss and results is performed. The location and quantity of the manufacturing center are determined by changing the capacity loss rate of the unreliable manufacturing center. The total cost, transportation cost, transportation cost, and transshipment amount in the objective function are shown in Table 3.

In the case of overcapacity, the manufacturing center will face the loss of processing capacity. In the face of customer demand uncertainty, the grey theory is used to change the uncertainty problem into the uncertainty probability problem, the transshipment strategy is used to supplement the capacity of unreliable manufacturing center, a three-level supply chain model is established to simulate the market

Table 3: Sensitivity analysis of capacity loss.

Number	Loss of capacity	Total cost	Central cost	Transportation cost	Transshipment cost	YR_i	YU_i
1	0.1	138071.3	123134.6	14924.04	12.66	4	1, 2, 3
2	0.15	138090.2	123134.6	14924.04	31.56	4	1, 2, 3
3	0.2	137975.5	123134.6	14782.23	58.67	4	1, 2, 3
4	0.25	138011	123134.6	14782.23	94.17	4	1, 2, 3
5	0.3	138054.8	123134.6	14782.23	137.97	4	1, 2, 3
6	0.35	138097.7	123134.6	14782.23	180.87	4	1, 2, 3
7	0.4	176405.9	162413	13752.93	239.97	2, 4	1, 3
8	0.45	176453.5	162413	13752.93	287.57	2, 4	1, 3
9	0.5	176241.5	162413	13467.33	361.17	2, 4	1, 3
10	0.55	176130.7	162413	13302.41	415.29	2, 4	1, 3
11	0.6	182986.7	169158.2	13302.41	526.1317	1, 2	3, 4
12	0.65	216218.8	201436.6	14782.2	0	1, 2, 3	—
13	0.7	216218.8	201436.6	14782.2	0	1, 2, 3	—
14	0.75	216218.8	201436.6	14782.2	0	1, 2, 3	—
15	0.8	216218.8	201436.6	14782.2	0	1, 2, 3	—
16	0.85	216218.8	201436.6	14782.2	0	1, 2, 3	—
17	0.9	216218.8	201436.6	14782.2	0	1, 2, 3	—
18	0.95	216218.8	201436.6	14782.2	0	1, 2, 3	—
19	1	216218.8	201436.6	14782.2	0	1, 2, 3	—

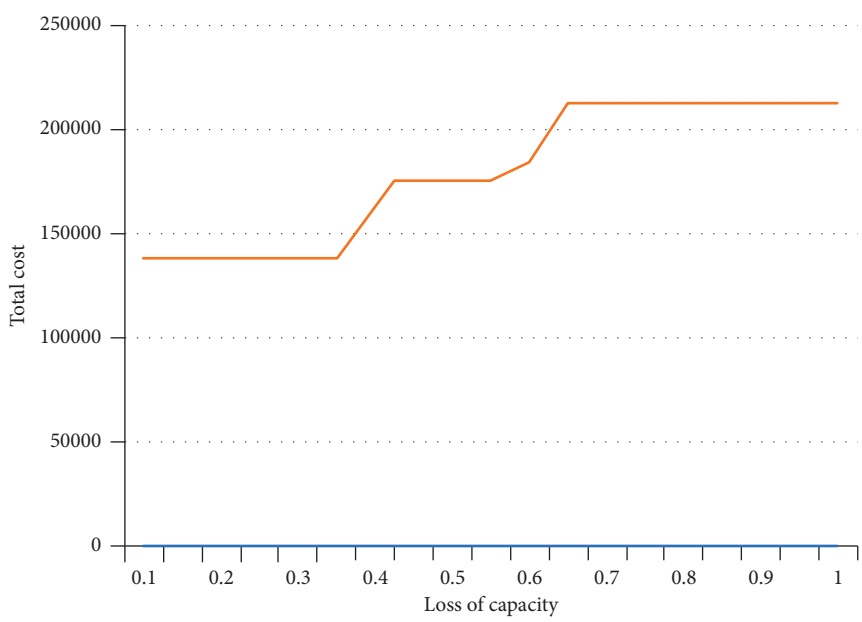


Figure 4: Total cost versus loss of capacity.

operation rules, and the sensitivity analysis of capacity loss to uncertain market demand is conducted. It is found that the greater the capacity loss of unreliable manufacturing center is, it is impossible that the total cost of supply chain and transportation shows an upward trend due to the insufficient operating rate of manufacturing centers. When the capacity loss reaches a certain degree, the manufacturing center with the lowest production capacity will be eliminated, and the overall cost is on the rise.

From the numerical results in columns 2 and 3 of Table 3, it can be concluded that when the capacity loss rate of modern industrial enterprises increases, the total cost of the supply chain shows an upward trend, and the manufacturing center also changes to a reliable manufacturing center (as shown in Figure 4). When the capacity loss reaches 0.65 and above, all manufacturing centers would have been converted into reliable manufacturing centers, and production centers with backward production capacity would have been discontinued.

In order to form the core competitiveness of manufacturing enterprises, it is suggested to develop reliable manufacturing centers as far as possible to cope with the uncertainty of market demand. If the market-oriented OEM manufacturing enterprises can save the construction and management costs of manufacturing centers, the stability of product manufacturing cannot be guaranteed, which may

cause supply interruption. Therefore, manufacturing enterprises need to build reliable manufacturing centers to ensure the flexibility of the supply chain.

4. Conclusions

This paper studies the problem of supply chain disruption in the downstream chain of manufacturing enterprises due to product foundry. The deep reason is that, in the face of risk decision-making, manufacturers will make decisions according to their own interests. This conflict between the manufacturer's interests and the supply chain's interests leads to the risk of disruption, which is more common in daily supply chain operations. When manufacturing companies face the risk of supply disruption, they will not only cause lost production capacity but also even lead to continuous business interruption. In order to solve the downstream supply disruption crisis of manufacturers under uncertain demand, this paper analyses the market elimination mechanism under the condition of product OEM production by constructing the global supply chain of three-level manufacturing enterprises and gives the supply chain elasticity optimization response strategy of manufacturing enterprises. The study found that the development of reliable manufacturing centers can effectively deal with the uncertainty of market demand, while market-oriented product OEMs can save the construction and management costs of manufacturing centers, but the stability of product supply is not stable. To the guarantee, it is likely to cause the risk of supply disruption. Therefore, in order to form the core competitiveness of manufacturing enterprises, the manufacturing supply chain design needs to flexibly select reliable manufacturing centers and unreliable manufacturing centers, reduce the total cost of the supply chain, and ensure the stable operation of the supply chain. In the complex market competition environment, manufacturing enterprises should actively respond, including the supply chain design and the recovery strategy plan after the interruption, so as to form a flexible global supply chain with more risk resistance and improve market competitiveness.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] C. F. Durach, J. Kembro, and A. Wieland, "A new paradigm for systematic literature reviews in supply chain management," *Journal of Supply Chain Management*, vol. 53, no. 4, pp. 67–85, 2017.

[2] G. Baryannis, S. Validi, S. Dani, and G. Antoniou, "Supply chain risk management and artificial intelligence: state of the art and future research directions," *International Journal of Production Research*, vol. 57, no. 7, pp. 2179–2202, 2019.

[3] T. J. Pettit, K. L. Croxton, and J. Fiksel, "The evolution of resilience in supply chain management: a retrospective on ensuring supply chain resilience," *Journal of Business Logistics*, vol. 40, no. 1, pp. 56–65, 2019.

[4] Y. Fan and M. Stevenson, "A review of supply chain risk management: definition, theory, and research agenda," *International Journal of Physical Distribution & Logistics Management*, vol. 48, no. 3, pp. 205–230, 2018.

[5] S. Hosseini, N. Morshedlou, D. Ivanov, M. D. Sarder, K. Barker, and A. A. Khaled, "Resilient supplier selection and optimal order allocation under disruption risks," *International Journal of Production Economics*, vol. 213, pp. 124–137, 2019.

[6] R. Dubey, A. Gunasekaran, S. J. Childe, T. Papadopoulos, C. Blome, and Z. Luo, "Antecedents of resilient supply chains: an empirical study," *Institute of Electrical and Electronics Engineers Transactions on Engineering Management*, vol. 66, no. 1, pp. 8–19, 2017.

[7] E. Durmić, Ž. Stević, P. Chatterjee, M. Vasiljević, and M. Tomašević, "Sustainable supplier selection using combined FUCOM–Rough SAW model," *Reports in Mechanical Engineering*, vol. 1, no. 1, pp. 34–43, 2020.

[8] A. Yousefi and M. S. Pishvaee, "A fuzzy optimization approach to integration of physical and financial flows in a global supply chain under exchange rate uncertainty," *International Journal of Fuzzy Systems*, vol. 20, no. 8, pp. 2415–2439, 2018.

[9] E. Rosca, G. Möllering, A. Rijal, and J. C. Bendul, "Supply chain inclusion in base of the pyramid markets," *International Journal of Physical Distribution & Logistics Management*, vol. 49, no. 5, pp. 575–598, 2019.

[10] S. Chakraborty, R. Chattopadhyay, and S. Chakraborty, "An integrated D-MARCOS method for supplier selection in an iron and steel industry," *Decision Making: Applications in Management and Engineering*, vol. 3, no. 2, pp. 49–69, 2020.

[11] B. Zahiri, J. Zhuang, and M. Mohammadi, "Toward an integrated sustainable-resilient supply chain: a pharmaceutical case study," *Transportation Research Part E: Logistics and Transportation Review*, vol. 103, pp. 109–142, 2017.

[12] N. Foroozesh, R. Tavakkoli-Moghaddam, S. M. Mousavi, and B. Vahdani, "A new comprehensive possibilistic group decision approach for resilient supplier selection with mean-variance-skewness-kurtosis and asymmetric information under interval-valued fuzzy uncertainty," *Neural Computing and Applications*, vol. 31, no. 11, pp. 6959–6979, 2019.

[13] A. Haeri, S. M. Hosseini-Motlagh, M. R. Ghatreh Samani, and M. Rezaei, "A mixed resilient-efficient approach toward blood supply chain network design," *International Transactions in Operational Research*, vol. 27, no. 4, pp. 1962–2001, 2020.

[14] D. Pamucar, M. Yazdani, R. Obradovic, A. Kumar, and M. Torres-Jiménez, "A novel fuzzy hybrid neutrosophic decision-making approach for the resilient supplier selection problem," *International Journal of Intelligent Systems*, vol. 35, no. 12, pp. 1934–1986, 2020.

[15] H. Golpîra, "Optimal integration of the facility location problem into the multi-project multi-supplier multi-resource construction supply chain network design under the vendor managed inventory strategy," *Expert Systems with Applications*, vol. 139, p. 112841, 2020.

[16] B. Vishkaei, S. T. A. Niaki, M. Farhangi, and I. Mahdavi, "A single-retailer multi-supplier multi-product inventory model with destructive testing acceptance sampling and inflation,"

Journal of Industrial and Production Engineering, vol. 36, no. 6, pp. 351–361, 2019.

[17] K. J. Mizgier, "Global sensitivity analysis and aggregation of risk in multi-product supply chain networks," *International Journal of Production Research*, vol. 55, no. 1, pp. 130–144, 2017.

[18] J. Liu, H. Zhou, and J. Wang, "The coordination mechanisms of emergency inventory model under supply disruptions," *Soft Computing*, vol. 22, no. 16, pp. 5479–5489, 2018.

[19] L. Coleman, "The power of resilience Yossi Sheffi the MIT press, Cambridge MA, 2015, 14 pp," *Journal of Contingencies and Crisis Management*, vol. 25, no. 2, pp. 114-115, 2017.

[20] B. Niu, J. Li, J. Zhang, H. K. Cheng, and Y. R. Tan, "Strategic analysis of dual sourcing and dual channel with an unreliable alternative supplier," *Production and Operations Management*, vol. 28, no. 3, pp. 570–587, 2019.

[21] C. Woarawichai and T. Naenna, "Solving inventory lot-sizing with supplier selection under alternative quantity discounts and vehicle capacity," *International Journal of Logistics Systems and Management*, vol. 30, no. 2, pp. 179–194, 2018.

[22] R. Pellegrino, N. Costantino, and D. Tauro, "Supply chain finance: a supply chain-oriented perspective to mitigate commodity risk and pricing volatility," *Journal of Purchasing and Supply Management*, vol. 25, no. 2, pp. 118–133, 2019.

[23] A. Qazi, A. Dickson, J. Quigley, and B. Gaudenzi, "Supply chain risk network management: a Bayesian belief network and expected utility based approach for managing supply chain risks," *International Journal of Production Economics*, vol. 196, pp. 24–42, 2018.

[24] X. Li, J. Chen, and X. Ai, "Contract design in a cross-sales supply chain with demand information asymmetry," *European Journal of Operational Research*, vol. 275, no. 3, pp. 939–956, 2019.

[25] J. Namdar, X. Li, R. Sawhney, and N. Pradhan, "Supply chain resilience for single and multiple sourcing in the presence of disruption risks," *International Journal of Production Research*, vol. 56, no. 6, pp. 2339–2360, 2018.

[26] Y. Han, W. K. Chong, and D. Li, "A systematic literature review of the capabilities and performance metrics of supply chain resilience," *International Journal of Production Research*, vol. 58, no. 15, pp. 4541–4566, 2020.

[27] S. DuHadway, S. Carnovale, and B. Hazen, "Understanding risk management for intentional supply chain disruptions: risk detection, risk mitigation, and risk recovery," *Annals of Operations Research*, vol. 283, no. 1-2, pp. 179–198, 2019.

[28] H. Shen, Y. Liang, Z.-J. M. Shen, and C.-P. Teo, "Reliable flexibility design of supply chains via extended probabilistic expanders," *Production and Operations Management*, vol. 28, no. 3, pp. 700–720, 2019.

[29] J. G. Wang, "Research on supply chain location game considering interruption risk," *Computer Engineering and Applications*, vol. 53, no. 20, pp. 231–236, 2017.

[30] D. Simchi-Levi, H. Wang, and Y. Wei, "Increasing supply chain robustness through process flexibility and inventory," *Production and Operations Management*, vol. 27, no. 8, pp. 1476–1491, 2018.

[31] R. Sreedevi and H. Saranga, "Uncertainty and supply chain risk: the moderating role of supply chain flexibility in risk mitigation," *International Journal of Production Economics*, vol. 193, pp. 332–342, 2017.

[32] L. T. T. Doan, Y. Amer, S.-H. Lee, P. N. K. Phuc, and L. Q. Dat, "A comprehensive reverse supply chain model using an interactive fuzzy approach-a case study on the Vietnamese electronics industry," *Applied Mathematical Modelling*, vol. 76, pp. 87–108, 2019.

[33] L. Jin, R. Mesiar, and R. Yager, "Parameterized preference aggregation operators with improved adjustability," *International Journal of General Systems*, vol. 49, no. 8, pp. 1–13, 2020.

[34] M. Fattahi and K. Govindan, "A multi-stage stochastic program for the sustainable design of biofuel supply chain networks under biomass supply uncertainty and disruption risk: a real-life case study," *Transportation Research Part E: Logistics and Transportation Review*, vol. 118, pp. 534–567, 2018.

[35] H. Sun, Q. Y. Cui, and Y. F. Xue, "Research on uncertain demand location-path robust optimization from the perspective of risk coping," *Operations Research and Management Science*, vol. 26, no. 11, pp. 26–34, 2017.

[36] B. Shen and Q. Li, "Market disruptions in supply chains: a review of operational models," *International Transactions in Operational Research*, vol. 24, no. 4, pp. 697–711, 2017.

[37] M. R. G. Samani and S.-M. Hosseini-Motlagh, "An enhanced procedure for managing blood supply chain under disruptions and uncertainties," *Annals of Operations Research*, vol. 283, no. 1-2, pp. 1413–1462, 2019.

[38] S. K. Paul, R. Sarker, and D. Essam, "A reactive mitigation approach for managing supply disruption in a three-tier supply chain," *Journal of Intelligent Manufacturing*, vol. 29, no. 7, pp. 1581–1597, 2018.

[39] L. M. Maiyar and J. J. Thakkar, "Robust optimisation of sustainable food grain transportation with uncertain supply and intentional disruptions," *International Journal of Production Research*, vol. 58, no. 18, pp. 5651–5675, 2020.

[40] B. C. Giri and S. Dey, "Game theoretic models for a closed-loop supply chain with stochastic demand and backup supplier under dual channel recycling," *Decision Making: Applications in Management and Engineering*, vol. 3, no. 1, pp. 108–125, 2020.

[41] A. Qamar, M. A. Hall, D. Chicksand, and S. Collinson, "Quality and flexibility performance trade-offs between lean and agile manufacturing firms in the automotive industry," *Production Planning & Control*, vol. 31, no. 9, pp. 723–738, 2020.

[42] J. Hong, Y. Zhang, and M. Ding, "Sustainable supply chain management practices, supply chain dynamic capabilities, and enterprise performance," *Journal of Cleaner Production*, vol. 172, pp. 3508–3519, 2018.

[43] A. Shishodia, P. Verma, and V. Dixit, "Supplier evaluation for resilient project driven supply chain," *Computers & Industrial Engineering*, vol. 129, pp. 465–478, 2019.

[44] M. Yazdani, C. Kahraman, P. Zarate, and S. C. Onar, "A fuzzy multi attribute decision framework with integration of QFD and grey relational analysis," *Expert Systems with Applications*, vol. 115, pp. 474–485, 2019.

[45] Z. Rezaee, "Supply chain management and business sustainability synergy: a theoretical and integrated perspective," *Sustainability*, vol. 10, no. 1, p. 275, 2018.

[46] B. Liu and K. Iwamura, "Chance constrained programming with fuzzy parameters," *Fuzzy Sets and Systems*, vol. 94, no. 2, pp. 227–237, 1998.

[47] B. Kim and K. S. Park, "Organizational structure of a global supply chain in the presence of a gray market: information asymmetry and valuation difference," *International Journal of Production Economics*, vol. 175, pp. 71–80, 2016.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 6672178, 9 pages
https://doi.org/10.1155/2021/6672178

Research Article

Application of the Fuzzy Optimal Model in the Selection of the Startup Hub

Xinman Zhu,[1] **Jie Dai**,[2] **Haoran Wei**,[2] **Debing Yang**,[3] **Weilun Huang**,[4] **and Zhang Yu**[5]

[1]*School of Management, Wenzhou Business College, Wenzhou, China*
[2]*Institute for Advanced Studies, University of Malaya, Kuala Lumpur, Malaysia*
[3]*School of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, China*
[4]*School of Finance and Trade, Wenzhou Business College, Wenzhou, China*
[5]*Department of Media and Communications, London School of Economics and Political Science, Holborn, UK*

Correspondence should be addressed to Weilun Huang; drweilunhuang@126.com

Received 15 November 2020; Revised 25 December 2020; Accepted 13 January 2021; Published 4 February 2021

Academic Editor: Jorge E. Macias-Diaz

This paper integrates nominal group technique (NGT), analytical hierarchy process (AHP), and fuzzy technique for order preference by similarity to an ideal solution (TOPSIS) approach, and a case study has been used to demonstrate the fuzzy optimal selection model. From a literature review on the startup hub and the interviews conducted with officials and experts, the selection criteria are (1) convenience—promoted by the city's entrepreneurial policies or its traffic infrastructure; (2) potentiality—promoted by a regional network or value chain of startups. Lastly, the best idle land resulted in this case study with equal decision-making power using the fuzzy method is Taipei Jianguo Brewery, and the difference of decision-making power might make the best idle land to be Wanbao Textile Factory.

1. Introduction

Many cities around the world are actively promoting various aspects of policies of innovation and entrepreneurship. Many works of the literature have proved that the startup hub could help founding startups, launching the successful venture, and upgrading urban economic and political status. Also, many startup hubs help startups in the country get economic or noneconomic assistance from the government, so startup hub policies are an essential part of innovation and entrepreneurship. Based on the literature review, the reason why startup hubs could help startups succeed is as follows: the competitive power of startups might be upgraded by the space and resources, as well as the interactive relations (such as information and communications technology, experts, and customers) in specialized services of the startup hub. Mulas et al. [1] found the entrepreneurship ecosystems (as startup hubs) in cities could maintain competitiveness, reduce poverty, and increase shared prosperity.

To increase competitive advantages of its startups, through startup hubs, Chongqing municipal government provides a variety of subsidies, such as the subsidy of production (up to RMB a million dollars), the subsidy proportion of rent (up to 50% for 200 square meters), the grant of research and development (RMB 0.5 million dollars, the maximum of a single project), the subsidy of patent maintenance fee (the full amount in three years), the subsidy of enterprise financing (up to RMB a million dollars), the equity investment of the maker (up to RMB a million dollars), and the startup project support fund (up to RMB two million dollars). Kupp et al. [2], Ledford [3], Katz and Wagner [4], and Bøllingtoft and Ulhøi [5] believed that the functions or values of the startup hub for startups are the space and resource sharing to make cost reduction, the provision of high-impact professional services and capital to make market-entry time reduction, the link of the manufacturer's cooperation network to establish the industry chain, and the alliance or integration on the external networks of startup hubs to accelerate the startup speed.

Moreover, the economic benefits offered by a startup hub for startups are demand-side economy of scale, the reduction of transaction cost, the spillover effect of spatial knowledge, urban innovation, and collective learning. The demand-side economy of scale can expand the scale of demand and then broaden cooperation in fields such as agriculture, energy, clean energy, and innovation. Due to the regional culture, the mode of production, and the regulations of startup hubs, the reduction of transaction costs can reduce the uncertainty of interaction and cooperation among startups and the transaction cost of the opportunistic behavior or hitchhiker. Tacit knowledge is an intangible cultural resource in the region of startup hubs, which could be thought as its unique spatial knowledge, and its spillover effect could reduce the cost or increase the benefit of startups. Urban innovation and collective learning are the communication and learning networks with the government, participating institutions, or informality, which are established by startup hubs [6].

The economic benefits of startup hubs to cities are their increment upgrading on residents' income, employment, and innovation environment. Frost-Kumpf [7] believed that the advantages of cultural districts, which are now thought to be a startup hub, for the city are as follows: (1) the beautification and activation of urban sightseeing and urban innovation environment; (2) the attraction of employment opportunities, residents, and high knowledge of labor; and (3) the value enhancement of the real estate taxation and mortgage. Also, the startup hubs could improve the status of the city in the global value chains through the interaction of startups, trade, and networks. The points of startup hubs should be the creation of new technologies, products, and markets caused by the integration of many kinds of technologies and industries, such as life sciences, telecommunications, energy, and education [8]. Katz and Wagner [4] believed that the startup hubs in the city could boost economic benefit, status, and jobs and reduce its carbon emissions.

For the advantages of startup hubs, many cities are eager to find suitable idle land which could be converted into startup hubs. From a review of the literature on startup hubs and the result of interviews conducted with experts, the considerations of suitable idle land that could be converted into startup hubs should be its site and architecture. Thus, this paper focuses on the selection process and empirical results on the site of startup hubs across the region, which are fostered by the municipal and local government. People in different countries have different ways of selecting the site of startup hubs. There are two kinds of decision-making party for selecting the site of startup hubs; one party is the private operator of the startup hubs, and the government only needs to make sure that the competition on selecting the site of startup hubs is engaged in a fair and open process, such as the case of the Silicon Valley in the United States. The other is the municipal government and the government-led integration of the space and resources required by startup hubs, such as the Chongqing case in China and the Paris case in France.

This paper cannot find any literature that discussed the site selection of startup hubs fostered by the municipal government. The decision-making process resulting in selecting the best site of idle lands that could be converted into startup hubs is complex, dynamic, and uncertain. Fuzzy TOPSIS should apply it to solve multiple-criteria decision-making problems in an uncertain environment. Therefore, the above decision-making process resulting in the selection of the best idle land in this paper is to use the nominal group technique (NGT) for qualitative and quantitative criteria, the concept of analytical hierarchy process (AHP) for selecting the best idle land, and fuzzy technique for order preference by similarity to an ideal solution (TOPSIS) method which could reduce the risk from these multiple fuzzy criteria.

The purpose of this paper is to propose and elaborate upon a multiple-criteria decision-making (MCDM) approach for the site selection of startup hubs fostered by the municipal government. From the results of the literature review and expert surveys, the selection criteria are its convenience and potentiality. As Loots et al. [9] thought, the ecosystem (like startup hub) has benefitted from social interactions and reciprocity transactions, and its growth and sustainability should be influenced by the city's local demand dynamics, competition for the physical space, gentrification, and policies [10–14].

The remainder of this paper is organized as follows: in Section 2, a summary of the review of the literature and the site selection process of startup hubs fostered by the municipal government are given. The results of a case study are presented in Section 3, focusing on the site selection of the best idle land, which could be converted into startup hubs by the Taipei municipal government of Taiwan. Section 4 concludes our findings and suggests further research.

2. Literature Review and the Proposed Method

Combining NGT and AHP with fuzzy TOPSIS provides an integrated process to solve the MCDM problem for site selection of startup hub candidates fostered or supported by the municipal government. The flowchart is shown in Figure 1 [15–20].

Firstly, from Figure 1, an initial list of idle lands in Taipei city that could be used as startup hubs should be compiled from the city government. In addition, startup hub candidates are chosen from the above initial list vetted with the experts who work in the relevant field in Taipei city.

Secondly, NGT was used to create the selection criteria of candidates based on a review of the relevant literature and results of key informant interviews conducted; the site selection decision of the best idle land which could be converted into startup hubs by the municipal government depends on the criteria on the convenience degree and the potential on its site of the idle land. The points of NGT are the join of every decision maker, no opinion leader, and all opinions are judged fairly. Thus, criteria can be divided into two types as convenience criteria C^C and potential criteria C^P. Then, the criteria of this paper are $C = \{C_1, C_2, \ldots, C_n\} = C^C \cup C^P$ and $C^C \cap C^P = \varnothing$ [21].

The convenience degree factor of startup hubs is its convenience, which is promoted by policies or traffic

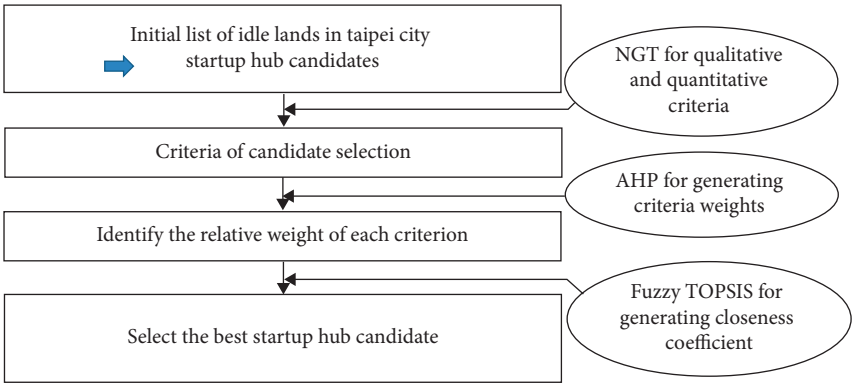


Figure 1: The integration process of site selection across a number of startup hub candidates.

infrastructure ($C^C = \{C_1, C_2\}$). In addition, the potential factor for the site is the potential which is promoted by the regional network and value chains of startups ($C^P = \{C_3, C_4, C_5\}$). Therefore, the government should clarify the present situation and consider the future of policy advice, traffic infrastructure, regional network, and value chains (production and marketing stages) of startups on the candidates' site. There are five criteria for the selection of the best candidates as the basis for startups.

Thirdly, the AHP is used to determine each criterion's relative weight based on preferences and subjective judgments of experts. Also, the fuzzy TOPSIS method is used to define the closeness coefficient for the performance of startup hub candidates with respect to each criterion. Chen et al. [22] applied MQCAC and fuzzy TOPSIS to improve the multicriteria decision analysis and the quality of unleaded gasoline.

Last but not least, the fuzzy TOPSIS approach could help to convert the decision makers' preferences and experience into meaningful results for decision makers by applying linguistic values to assess each alternative candidate based on each criterion [23]. Then, the best startup hub has been selected as a candidate.

To explore the second step of site selection's integration process, its optimal site selection criteria of the startup hub should depend on its theoretical and empirical literature. The resource accessibility of the startup hub would influence its competitive advantage and economic performance. The resource-based theory should be able to explain, predict, and examine the phenomena of the startup hub; its points are resource heterogeneity, entrepreneurial cognition, entrepreneurial alertness, market opportunity, coordinated knowledge and firm, ex-post and ex-ante limits to competition, uncertainty, information asymmetries, imperfect factor mobility, and path dependency [24, 25]. On the contrary, a thriving startup hub should have many interactions among startups and their surrounding political, socioeconomic, traffic, industrial, and local cultural environments [13, 14, 26–29].

From the above discussion, the reasons for the criteria of its convenience promoted by the government policy are as follows: (1) the convenience of startup hubs for startups could be promoted by a variety of policies, such as the grand funding, the preferential tax, and the construction of surrounding public facilities, as well as the system of land acquisition. (2) There are more and more policies related to startup hubs, such as supporting them in the innovation policy, adopting a new regional development policy, and applying the industrial policy of a country, as well as policies of small and medium-sized enterprises (SMEs). Also, more policies are related to startup hubs and are more convenient for startups. Brown and Mason [30] thought the policies for startup hubs should be designed to help foster vertical connections across ecosystems. Migendt et al. [31] found the finance-innovation-policy nexus existed.

The reasons for the criteria of its convenience promoted by traffic infrastructures are as follows: (1) the convenience for startups that startup hubs offer could be promoted by many kinds of traffic infrastructures, such as airport, railway station, and bus station, as well as mass rapid transit (MRT) station. (2) Most of these startups come from SMEs and need to make more of an effort to negotiate with customers or develop a new go-to-market field. Therefore, more and more traffic infrastructures around startup hubs in Taipei city make the startups more convenient. Ustinovichius et al. [32] thought the accessibility, transportation costs, and time spent on the road are important for improving business conditions. Katz and Wagner [4] found North Carolina's Research Triangle Park, in November 2012, unveiled a new 50-year master plan that calls for a light rail transit line to connect with the Raleigh-Durham region. Sorenson and Stuart [33] found the geographic and social networks between venture capitals and the startup would affect the identification or support of the startups' performance by venture capitals.

The reasons for the criteria of its potential promoted by the regional network are as follows:

(1) The regional network about education, culture, and service facilities could promote the potential of startup hubs, such as many works of the literature proved the essential factors for the potential of startup hubs which are as follows: A. knowledge spillover effects and spatial interactions on the output of universities and research institutions; B. the culture value of famous landmarks and

cultural landscapes; C. the living function of convenience stores; D. a large variety of leisure and entertainment places (see [4, 34–36]; [6, 37–41]).

(2) Various startup business programs are mostly related to the cultural and creative industries, for the potential—promoted by the regional network of startup hubs, and many governments would combine startup hubs with cultural and creative facilities [7, 42]. As Singh and Bala Subrahmanya [43] found, tech startups' financial requirements over their life cycle are their human capital, research capital, and social capital.

The reasons for the criteria of its potential promoted by startups' value chains which were divided with the stages of production and marketing are as follows:

(1) According to Webber and Labaste [44], Kaplinsky and Morris [45], and Chesbrough and Rosenbloom [46], all stages of startups' value chain are production, processing, brand, marketing, as well as logistics, and cross-industry cooperation. Ngugi et al. [47] found that the winning strategic factors responsible for mobile product innovation in the African market should be sustainability and reconfiguration of innovation, strategic assets, and business ecosystem. Geibel and Manickam [48], Herrmann et al. [49], Isenberg [50], and Belussi [6] stated that the startup ecosystem which is similar to a startup value chain would promote the potential of the startup hubs, for startups often lack relationships among customers, suppliers, and related industries. Kritikos [8] found that Berlin-Adlershof research hub is an integrated business, technology, and media park that enjoys a site's advantage by its easy access to scientific research and Berlin's commercial infrastructure.

The members of the startup ecosystem are colleges, startups, accelerators, angel investors (or seed fund), venture capitalists (or crowd-funding), mergers and acquisitions in the banking industry, business partners, and service providers, as well as governments and customers. Vankamamidi [51] found that T-Hub was formed to build a booming startup ecosystem in the country. Baum and Silverman [52] found that alliance, intellectual, and human capital of biotechnology startups are the selection criteria in venture financing and the affecting factors on biotechnology startups' future performance. Hellmann and Puri [53] found venture capitalists could provide financial resources or management expertise to startups. Baum et al. [54] found that the variation in startups' alliance network composition supports their early performance, especially their innovative performance. Walker et al. [55] found strong support for social capital in reproducing the startups' work overtime.

(2) Startup hubs promoted by the governments have the potential to become the influential startup ecosystem, such as Xiantao Big Data Valley of Chongqing in China and Welcome City Lab of Paris in France. The members of the startup ecosystem in Xiantao Big Data Valley include more than 30 leading companies, more than 300 core companies, more than 1000 startups, more than 200 incubations, more than 500 memberships, more than 20 research and development institutions, and more than ten training systems. The members of the startup ecosystem in Welcome City Lab are the incubations, startups, and partners, such as Air France, Amadeus company, Paris Aéroport, and Carlson Wagonlit Travel (CWT), as well as Galeries Lafayette Group [56].

3. A Case Study of the Site Selection of the Best Idle Land Which Could Be Converted into Startup Hubs by the Municipal Government

The decision-making group (DMG, seven experts) consists of a government officer (D_1), two experts of the startup hub (D_2 and D_3), and four startups (D_4, D_5, D_6, and D_7) who are invited to participate in providing a range of opinions based on their judgment. With the review of the Taipei municipal government, this paper obtains an initial list of more than ten idle lands. The Taipei municipal government would like to select the best idle land which could be converted into startup hubs, and hub candidates resulted from the initial list of idle lands vetted by the DMG are Jianguo Brewery (V_1), Wanbao Textile Factory (V_2), and Minglun Elementary School (V_3).

To select the best idle land which could be converted into startup hubs, the candidates could be determined as follows: (1) C_1 (policy): its convenience promoted by government's policy, (2) C_2 (traffic infrastructure): its convenience promoted by traffic infrastructure, (3) C_3 (regional network): its potential promoted by the regional network, (4) C_4 (value chain (production stages)): its potential promoted by the production stages of startups' value chain, and (5)C_5 (value chain (marketing stages)): its potential promoted by the marketing stages of startups' value chain. The hierarchical and sequential structure of group decision-making is presented in Figure 2.

As an illustrative example, five experts presented their points of view for site selection by linguistic variables to assess the weight that reflects the importance of criteria, and the matrix of triangular fuzzy numbers and fuzzy preference used in experts for linguistic variables [very low; low; fairly low; fairly high; high; very high] are [(0, 0, 0.2); (0, 0.2, 0.4); (0.2, 0.4, 0.6); (0.4, 0.6, 0.8); (0.6, 0.8, 1.0); (0.8, 1.0, 1.0)]. Moreover, the linguistic variables of fuzzy ratings for the alternative's preference are from Figure 3. In Figure 3, a triangular fuzzy number can be defined as a triplet (α, β, γ); the utility function $(u_\delta(x))$ of the fuzzy rating δ is defined as [57]

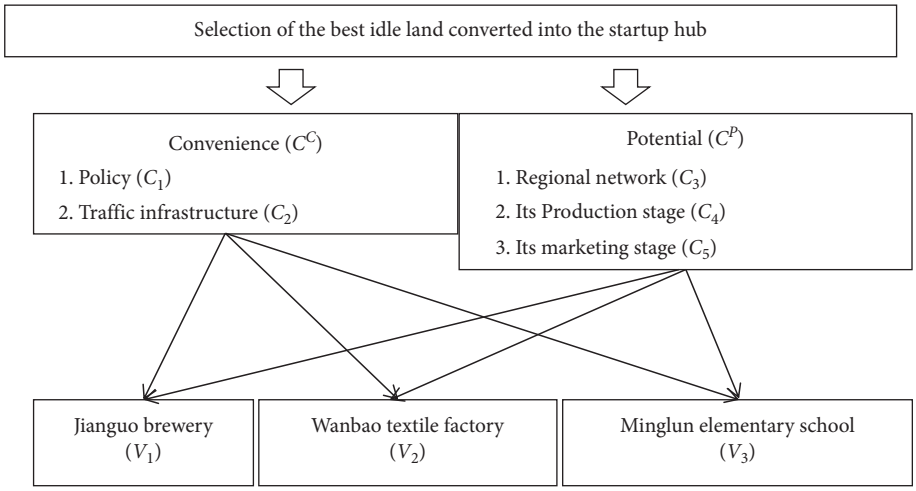


Figure 2: Hierarchy structure of the selection for the startup hub.

$$u_\delta(x) = \begin{cases} \dfrac{(x-\alpha)}{(\beta-\alpha)}, & x \in [\alpha, \beta], \\ \dfrac{(\gamma-x)}{(\gamma-\beta)}, & x \in [\beta, \gamma], \\ 0, & \text{otherwise}, \end{cases} \quad (1)$$

where $\alpha \le \beta \le \gamma$.

The fuzzy weights of evaluation criteria generally determined by the seven experts are shown in Table 1. In addition, the seven experts use linguistic variables in order to evaluate the fuzzy preference rating of each candidate with respect to each criterion. The rating is shown in Table 2. Furthermore, the linguistic variables for the importance of each criterion's fuzzy weights are similar to Figure 3.

So far, most works of the literature on the MCDM problem have assumed the weight of each decision maker is the same in their selection model. In a practical case, this is rarely true. By referring to the site selection of the startup hub as an example, the government will most likely have the most authority and influence power in the decision. Therefore, two selection models have been discussed in this paper. Following most historical studies, case one assumes that the weight of decision-making power for each decision maker is equal. Case two considers that the government has half of the decision-making power, and then, the rest of the decision makers equally shared the remaining half.

SMEs' innovation and industrial policies for all candidates are the same. The regional development policies are different for Wanbao Textile Factory and Minglun Elementary School. However, from a regional development perspective, there is no single regional development policy for Jianguo Brewery. The time-efficient (by car or on foot) orders from startup hub candidates to traffic infrastructures are Minglun Elementary School, Wanbao Textile Factory, and Jianguo Brewery (from the biggest to the smallest). Based on the number of education services, culture and recreation, and service facilities within 1 kilometer of candidates, the candidates' order on their regional network's promoted potential is Jianguo Brewery, Wanbao Textile Factory, and Minglun Elementary School (from the biggest to the smallest). For Jianguo Brewery, other value chains for the tourism industry, digital industry, creative industry, and the neighboring Taipei area are quite far away. For Wanbao Textile Factory and Minglun Elementary School, the tourism supply chain exists in the neighboring area; however, there are no startup's auxiliary industries, such as intellectual property, law, and accounting. The fuzzy preference structure is shown in Table 2.

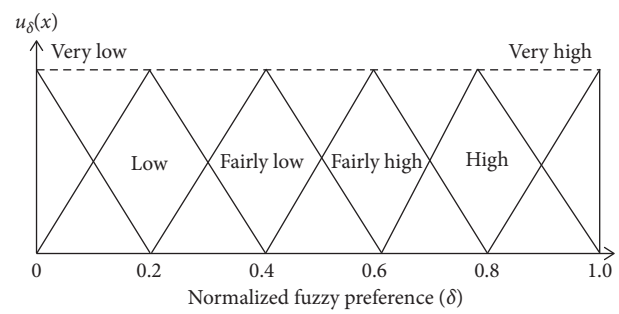


Figure 3: The linguistic variables of fuzzy ratings for the alternative's preference.

The weights of each criterion summarized in Table 1 and the fair consistency evaluation in fuzzy preference relations for policy subjects by experts summarized in Table 2 are used to create a fuzzy weighted decision matrix, which is shown in Table 3. The data of Table 3 are from $V^* = [(\max_i \tilde{v}_{ij} | j \in J), (\min_i \tilde{v}_{ij} | j \in J)]$ and $V^- = [(\max_i \tilde{v}_{ij} | j \in J), (\min_i \tilde{v}_{ij} | j \in J)]$, and the fuzzy positive ideal and fuzzy negative ideal are determined as

$$\begin{aligned} V^* &= [(1,1,1), (1,1,1), (1,1,1), (0.04, 0.04, 0.04), (1,1,1)], \\ V^- &= [(0.16, 0.16, 0.16), (0.16, 0.16, 0.16), (0.12, 0.12, 0.12), \\ &\quad \cdot (0.8, 0.8, 0.8), (0.08, 0.08, 0.08)]. \end{aligned} \quad (2)$$

Table 1: The fuzzy weight of criteria from seven experts.

	C_1	C_2	C_3	C_4	C_5
D_1	(0.8, 1.0, 1.0)	(0.2, 0.4, 0.6)	(0.6, 0.8, 1.0)	(0.4, 0.6, 0.8)	(0.4, 0.6, 0.8)
D_2	(0.6, 0.8, 1.0)	(0.4, 0.6, 0.8)	(0.4, 0.6, 0.8)	(0.4, 0.6, 0.8)	(0.8, 1.0, 1.0)
D_3	(0.6, 0.8, 1.0)	(0.4, 0.6, 0.8)	(0.4, 0.6, 0.8)	(0.4, 0.6, 0.8)	(0.8, 1.0, 1.0)
D_4	(0.4, 0.6, 0.8)	(0.4, 0.6, 0.8)	(0.6, 0.8, 1.0)	(0.6, 0.8, 1.0)	(0.6, 0.8, 1.0)
D_5	(0.4, 0.6, 0.8)	(0.4, 0.6, 0.8)	(0.6, 0.8, 1.0)	(0.6, 0.8, 1.0)	(0.6, 0.8, 1.0)
D_6	(0.6, 0.8, 1.0)	(0.2, 0.4, 0.6)	(0.6, 0.8, 1.0)	(0.4, 0.6, 0.8)	(0.8, 1.0, 1.0)
D_7	(0.6, 0.8, 1.0)	(0.2, 0.4, 0.6)	(0.6, 0.8, 1.0)	(0.6, 0.8, 1.0)	(0.8, 1.0, 1.0)
Fuzzy weight (case 1)	(0.40, 0.76, 1.00)	(0.20, 0.50, 0.80)	(0.40, 0.74, 1.00)	(0.40, 0.68, 1.00)	(0.40, 0.87, 1.00)
Fuzzy weight (case 2)	(0.40, 0.85, 1.00)	(0.20, 0.46, 0.80)	(0.40, 0.76, 1.00)	(0.40, 0.64, 1.00)	(0.40, 0.75, 1.00)

Table 2: Fuzzy preferences for the three candidates by seven experts using various criteria.

	F_1	F_2	F_3	F_4	F_5
V_1	(0.6, 0.8, 1.0)	(0.8, 1.0, 1.0)	(0.6, 0.8, 1.0)	(0.8, 1.0, 1.0)	(0.8, 1.0, 1.0)
V_2	(0.8, 1.0, 1.0)	(0.6, 0.8, 1.0)	(0.6, 0.8, 1.0)	(0.4, 0.6, 0.8)	(0.4, 0.6, 0.8)
V_3	(0.6, 0.8, 1.0)	(0.4, 0.6, 0.8)	(0.8, 1.0, 1.0)	(0.4, 0.6, 0.8)	(0.4, 0.6, 0.8)

Table 3: Fuzzy weighted decision matrix.

		F_1	F_2	F_3	F_4	F_5
Case 1	V_1	(0.24, 0.61, 1.00)	(0.16, 0.50, 0.80)	(0.24, 0.59, 1.00)	(0.32, 0.68, 1.00)	(0.32, 0.87, 1.00)
	V_2	(0.32, 0.76, 1.00)	(0.12, 0.40, 0.80)	(0.24, 0.59, 1.00)	(0.16, 0.41, 0.80)	(0.16, 0.52, 0.80)
	V_3	(0.24, 0.61, 1.00)	(0.08, 0.30, 0.64)	(0.32, 0.74, 1.00)	(0.16, 0.41, 0.80)	(0.16, 0.52, 0.80)
Case 2	V_1	(0.24, 0.68, 1.00)	(0.16, 0.46, 0.80)	(0.24, 0.61, 1.00)	(0.32, 0.64, 1.00)	(0.32, 0.75, 1.00)
	V_2	(0.32, 0.85, 1.00)	(0.12, 0.37, 0.80)	(0.24, 0.61, 1.00)	(0.16, 0.39, 0.80)	(0.16, 0.45, 0.80)
	V_3	(0.24, 0.68, 1.00)	(0.08, 0.27, 0.64)	(0.32, 0.76, 1.00)	(0.16, 0.39, 0.80)	(0.16, 0.45, 0.80)

In two cases, the shortest distance from the fuzzy positive ideal solution (FPIS, $d(V_i, V^*)$, $i = 1, 2, 3$) and the farthest distance from the fuzzy negative ideal solution (FNIS, $d(V_i, V^-)$, $i = 1, 2, 3$) can be determined by using the method of Zhang et al. (2011). The distance of each candidate from FPIS and FNIS is calculated with respect to the criterion, respectively; using $d_i^* = \sum_{j=1}^{n} d(\tilde{v}_{ij}, \tilde{v}_j^+)$ and $d_i^- = \sum_{j=1}^{n} d(\tilde{v}_{ij}, \tilde{v}_j^-)$, the distance of each candidate's fuzzy positive ideal and fuzzy negative ideal can be calculated with respect to each criterion, respectively, as shown in Table 4.

In case 1, the closeness coefficient ($CC_i = d_i^- / (d_i^* + d_i^-)$, $i = 1, 2, 3$) for each candidate is calculated as $CC_1 = 0.502$, $CC_2 = 0.501$, and $CC_3 = 0.484$. The alternatives are ranked by CC_i^*, and the results are $V_1 > V_2 > V_3$. Therefore, some measures are based on quantitative criteria; the best idle land which could be converted into startup hubs, in this case, is Jianguo Brewery. In case 2, each candidate's closeness coefficient is calculated as $CC_1 = 0.500$, $CC_2 = 0.502$, and $CC_3 = 0.485$, and the results are $V_2 > V_1 > V_3$. Therefore, the difference of decision-making power might make the best idle land which could be converted into startup hubs to be Wanbao Textile Factory.

Table 4: Distance between V_i ($i = 1, 2, 3$) and V^* (V^-) with respect to each candidate.

		F_1	F_2	F_3	F_4	F_5
Case 1	$d(V_1, V^*)$	0.494	0.575	0.499	0.649	0.399
	$d(V_2, V^*)$	0.416	0.625	0.499	0.459	0.569
	$d(V_3, V^*)$	0.494	0.698	0.421	0.459	0.569
	$d(V_1, V^-)$	0.654	0.420	0.614	0.308	0.643
	$d(V_2, V^-)$	0.712	0.396	0.614	0.434	0.425
	$d(V_3, V^-)$	0.654	0.293	0.667	0.434	0.425
Case 2	$d(V_1, V^*)$	0.476	0.589	0.493	0.638	0.419
	$d(V_2, V^*)$	0.402	0.637	0.493	0.454	0.592
	$d(V_3, V^*)$	0.476	0.708	0.416	0.454	0.592
	$d(V_1, V^-)$	0.677	0.408	0.620	0.313	0.599
	$d(V_2, V^-)$	0.744	0.389	0.620	0.440	0.405
	$d(V_3, V^-)$	0.677	0.289	0.676	0.440	0.405

4. Conclusion

From the results of the literature review, expert survey, and the above discussion, this paper can come to the conclusion that the site selection of startup hubs is important for startups, the government, and researchers. Therefore, this

paper focuses on the decision-making of the optimal site selection of startup hubs fostered by the municipal government. There are higher levels of uncertainty and complexity for the decision-making process resulting in the selection of policy subjects. Therefore, how to choose the best startup hub from candidates (i.e., idle lands) in the Taipei municipal government is an MCDM problem.

There are lots of idle lands in Taipei city that could be converted into startup hubs, but there is little literature about case studies on the subjects of the hub site consideration. As there is little literature on this topic, further investigation would be needed. According to the relevant provisions of the domestic and foreign-related cases, the literature review, and the result of interviews conducted with experts, the criteria for the selection of site candidates for startup hubs are its convenience promoted by the government's policy or traffic infrastructure and its potential promoted by the regional network or value chains of startups.

The satisfied idle land which could be converted into startup hubs is found by the integrated approach of NGT, AHP, and fuzzy TOPSIS method to assist decision makers' selection according to their preferences and resource limitations. In the end, the best idle land resulted in this case study with equal decision-making power is Jianguo Brewery, and the difference of decision-making power might make the best idle land to be Wanbao Textile Factory. The potential (the regional network and value chains of startups) for startup hubs of Jianguo Brewery is excellent. Also, the convenience (policy and traffic infrastructure) for startup hubs of Wanbao Textile Factory is superior. In addition, for startup hubs of Minglun Elementary School, its evaluations of potential and convenience are best for it is in the medial position of the city.

More precisely, this paper's contribution and the managerial insights are as follows: (1) the supplement on the literature of a startup hub, which is its optimal site selection: this paper first finds out that the existing literature lacks details on the municipal government's startup hub site selection. Although some existing literature and documents are discussing related issues of startup hubs, there is little empirical literature on the investigation of criteria to choose the best idle land which could be converted into startup hubs. (2) The empirical practice of the method to select the best idle land which could be converted into the startup hub: based on the sample of Taipei's idle lands, this paper summarizes and confirms the site selection, further expands the empirical research in this field, and provides the method and result of Taipei's best idle land, which could be converted into the startup hub.

This paper also stresses the research limitations and provides a useful recommendation for future research. This paper also provides thinking that the site selection consideration of startup hubs is an MCDM problem. To deal with the issues of fuzzy evaluation process among criteria, such as its uncertain factors, its discrimination factors, and its feedback and internal dependency, the future research of the MCDM method based on fuzzy theory, ARAS, and MCGP would make the evaluation results more accurate. Finally, this paper provides the best idle land that could be converted into startup hubs for the Taipei municipal government.

As this paper suggested, the MCGP formulation can be defined as the following:

$$\begin{aligned}
&\text{Min} \quad \sum_{i=1}^{n} w_i\left[\left(d_i^+ + d_i^-\right) + \left(e_i^+ + e_i^-\right)\right] \\
&s.t. \quad f_i(X) - d_i^+ + d_i^- = g_{i,1} \text{ or } g_{i,2} \text{ or } g_{i,m}, \quad i = 1, 2, \ldots, n \\
&\qquad g_i - e_i^+ + e_i^- = g_{i,\max}, \quad i = 1, 2, \ldots, n \\
&\qquad g_{i,\max} \geq g_i \geq g_{i,\min}, \quad i = 1, 2, \ldots, n \\
&\qquad d_i^+, d_i^-, e_i^+, e_i^- \geq 0, \quad i = 1, 2, \ldots, n, \\
&\qquad X \in F\,(F \text{ is a feasible set}),
\end{aligned} \tag{3}$$

where g_{ij} ($i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$) is the jth aspiration level of the ith goal, w_i represents the weight attached to the deviation, and d_i is the deviation from the target value g_i; $d_i^+ = \max(0, f_i(x) - g_i)$ and $d_i^- = \max(0, g_i - f_i(x))$ denote over- and underachievement of the ith goal, respectively. In addition, e_i^+ and e_i^- are positive and negative deviations attached to $|g_i - g_{i,\max}|$; $g_{i,\min}$ and $g_{i,\max}$ are, respectively, lower and upper bounds of g_i.

Future research could focus on the reaction potential of large-scale emergency incidents, which should be included in startup hubs' site selection criteria, such as coronavirus (COVID-19). Rowan and Galanakis [58] thought COVID-19 might help cross-cutting disruption in agri-food, ICT, health, and environment. On the contrary, future research could study the impacts of startup hubs on the socioeconomic environment of the city. Rowan and Galanakis [58] thought that startup hubs have the potential to accelerate the socioeconomic recovery of the city.

Future research could focus on the optimal site selection process of the startup hub, which might be different in different industries, regions, or entrepreneurs. Malecki [14] thought the entrepreneurial ecosystem might be appropriate only for the region with a critical mass of startups. Spigel and Harrison [27] thought that entrepreneurs might be excluded from some local socioeconomic networks for their gender, race, age, or education level.

Data Availability

Requests for access to the data used to support the findings of this study should be made to Weilun Huang (drweilunhuang@126.com).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by the Research Project of Zhejiang Federation of Social Sciences of China (no.

2021N101), Wenzhou Philosophy and Social Science Planning Project (no. 20wsk095), and Basic Scientific Research Projects of Wenzhou Science and Technology Bureau in 2020 (no. R2020007).

References

[1] V. Mulas, M. Minges, and H. Applebaum, "Boosting tech innovation: ecosystems in cities: a framework for growth and sustainability of urban tech innovation ecosystems," *Innovations: Technology, Governance, Globalization*, vol. 11, no. 1-2, pp. 98–125, 2016.

[2] M. Kupp, M. Marval, and P. Borchers, "Corporate accelerators: fostering innovation while bringing together startups and large firms," *Journal of Business Strategy*, vol. 38, no. 6, pp. 47–53, 2017.

[3] H. Ledford, "Start-ups fight for a place in Boston's biotech hub," *Nature*, vol. 522, no. 7555, pp. 138-139, 2015.

[4] B. Katz and J. Wagner, "The rise of innovation districts: a new geography of innovation in america," 2014.

[5] A. Bøllingtoft and J. P. Ulhøi, "The networked business incubator-leveraging entrepreneurial agency?" *Journal of Business Venturing*, vol. 20, no. 2, pp. 265–290, 2005.

[6] F. Belussi, "Policies for the development of knowledge-intensive local production systems," *Cambridge Journal of Economics*, vol. 23, no. 6, pp. 729–747, 1999.

[7] H. A. Frost-Kumpf, *Cultural Districts: The Art as a Strategy for Revitalizing Our Cities*, Americans for the Arts, Washington, D. C., USA, 1998.

[8] A. S. Kritikos, *A Hub for Startups but not (yet) for Fast-Growing Companies*, German Institute for Economic Research, Berlin, Germany, 2016.

[9] E. Loots, M. Neiva, L. Carvalho, and M. Lavanga, "The entrepreneurial ecosystem of cultural and creative industries in Porto: a sub-ecosystem approach," *Growth and Change*, vol. 3, pp. 1–22, 2020.

[10] M. Lavanga, "Cultural districts," in *Handbook of Cultural Economics*, R. Towse and T. Navarrete, Eds., pp. 174–182, Edward Elgar Publishing, London, UK, 3rd edition, 2020.

[11] M. Lavanga, E. Loots, A. Konomi, V. Avdikos, and E. Iliopoulou, "Creative entrepreneurship and urban space: exploring the location preferences of creative professionals in athens during the economic recession," *The Greek Review of Social Research (Special Issue "Work in the Aftermath of the Global Economic Crisis: New Spaces and Types of Work Emerging in Cities of Recession")*, vol. 153, pp. 5–36, 2020.

[12] J. Merkel, ""Freelance isn't free." Co-working as a critical urban practice to cope with informality in creative labour markets," *Urban Studies*, vol. 56, no. 3, pp. 526–547, 2019.

[13] D. Audretsch, C. Mason, M. P. Miles, and A. O'Connor, "The dynamics of entrepreneurial ecosystems," *Entrepreneurship & Regional Development*, vol. 30, no. 3-4, pp. 471–474, 2018.

[14] E. J. Malecki, "Entrepreneurship and entrepreneurial ecosystems," *Geography Compass*, vol. 12, no. 3, p. e12359, 2018.

[15] Z. K. Öztürk and Ş. Toptanci, "An integrated MCDM model for occupational safety specialist selection," *Journal of Business Research-Turk*, vol. 9, no. 4, pp. 419–435, 2017.

[16] K.-H. Chen, C.-N. Liao, and L.-C. Wu, "A selection model to logistic centers based on TOPSIS and MCGP methods: the case of airline industry," *Journal of Applied Mathematics*, vol. 2014, p. 1, 2014.

[17] Y. F. Chang and H. ISHII, "Probing the implementation of project management office by using DEMATEL with a hybrid MCDM model," *Journal of Japan Society for Fuzzy Theory and Intelligent Informatics*, vol. 25, no. 6, pp. 935–948, 2013.

[18] X. Zhang, "Venture capital investment selection decision-making base on fuzzy theory," *Physics Procedia*, vol. 25, pp. 1369–1375, 2012.

[19] C.-N. Liao and H.-P. Kao, "Supplier selection model using taguchi loss function, analytical Hierarchy process and multi-choice goal programming," *Computers & Industrial Engineering*, vol. 58, no. 4, pp. 571–577, 2010.

[20] S. Önü and S. Soner, "Transshipment site selection using AHP and TOPSIS approaches under fuzzy environment," *Waste Management*, vol. 28, no. 9, pp. 1552–1559, 2008.

[21] C. N. Liao, "An evaluation model using fuzzy TOPSIS and goal programming for TQM consultant selection," *Journal of Testing and Evaluation*, vol. 41, no. 1, pp. 122–130, 2013.

[22] K.-S. Chen, C.-H. Hsu, L.-Y. Ouyang, and C.-M. Yang, "Applying MQCAC and fuzzy TOPSIS to improve the unleaded gasoline quality," *Journal of Testing and Evaluation*, vol. 45, no. 3, pp. 20150426–20151057, 2017.

[23] W. Huang and Q. Zhang, "Selecting the optimal economic crop in minority regions with the criteria about soil and water conservation," *Agricultural Water Management*, vol. 241, p. 106295, 2020.

[24] J. Barney and S. A. Alvarez, "Resource-based theory and the entrepreneurial firm," *Strategic Entrepreneurship: Creating a New Mindset*, vol. 2, pp. 87–105, 2017.

[25] R. S. Nason and J. Wiklund, "An assessment of resource-based theorizing on firm growth and suggestions for the future," *Journal of Management*, vol. 44, no. 1, pp. 1–29, 2015.

[26] E. Stam and A. van de Ven, "Entrepreneurial ecosystem elements," *Small Business Economics*, vol. 12, pp. 1–24, 2019.

[27] B. Spigel and R. Harrison, "Toward A process theory of entrepreneurial ecosystems," *Strategic Entrepreneurship Journal*, vol. 12, no. 1, pp. 151–168, 2018.

[28] Z. J. Acs, E. Stam, D. B. Audretsch, and A. O'Connor, "The lineages of the entrepreneurial ecosystem Approach," *Small Business Economics*, vol. 49, no. 1, pp. 1–10, 2017.

[29] J. Alvedalen and R. Boschma, "A critical review of entrepreneurial ecosystems research: towards A future research agenda," *European Planning Studies*, vol. 25, no. 6, pp. 887–903, 2017.

[30] R. Brown and C. Mason, "Looking inside the spiky bits: a critical review and conceptualisation of entrepreneurial ecosystems," *Small Business Economics*, vol. 49, no. 1, pp. 11–30, 2017.

[31] M. Migendt, F. Polzin, F. Schock, F. A. Täube, and P. von Flotow, "Beyond venture capital: an exploratory study of the finance-innovation-policy nexus in cleantech," *Industrial and Corporate Change*, vol. 26, no. 6, pp. 973–996, 2017.

[32] L. Ustinovichius, A. Komarovska, and R. Komarovski, "Methods of determining the region's investment strategy," *Procedia Engineering*, vol. 182, pp. 732–738, 2017.

[33] O. Komarovski and T. E. Stuart, "Syndication networks and the spatial distribution of venture capital investments," *American Journal of Sociology*, vol. 106, no. 6, pp. 1546–1588, 2001.

[34] R. Florida and C. Mellander, "Rise of the startup city," *California Management Review*, vol. 59, no. 1, pp. 14–38, 2016.

[35] J. Youtie and P. Shapira, "Building an innovation hub: a case study of the transformation of university roles in regional technological and economic development," *Research Policy*, vol. 37, no. 8, pp. 1188–1204, 2008.

[36] M. Gittelman, "Does geography matter for science-based firms? Epistemic communities and the geography of research

and patenting in biotechnology," *Organization Science*, vol. 18, no. 4, pp. 724–741, 2007.

[37] M. Fritsch and V. Slavtchev, "Universities and innovation in space," *Industry and Innovation*, vol. 14, no. 2, pp. 201–218, 2007.

[38] M. M. Fischer and A. Varga, "Spatial knowledge spillovers and university research: evidence from Austria," *The Annals of Regional Science*, vol. 37, no. 2, pp. 303–322, 2003.

[39] L. Anselin, A. Varga, and Z. Acs, "Local geographic spillovers between university research and high technology innovations," *Journal of Urban Economics*, vol. 42, no. 3, pp. 422–448, 1997.

[40] M. P. Feldman and R. Florida, "The geographic sources of innovation: technological infrastructure and product innovation in the United States," *Annals of the Association of American Geographers*, vol. 84, no. 2, pp. 210–229, 1994.

[41] A. Jaffe, "Real effect of academic research," *The American Economic Review*, vol. 79, no. 5, pp. 957–970, 1989.

[42] F. Moulaert and F. Sekia, "Territorial innovation models: a critical survey," *Regional Studies*, vol. 37, no. 3, pp. 289–302, 2003.

[43] S. Singh and M. H. Bala Subrahmanya, "The financial requirements of tech startups over its lifecycle in Bangalore: an analysis of why and how do they differ?" *International Journal of Finance & Economics*, vol. 2020, pp. 1–19, 2020.

[44] C. M. Webber and P. Labaste, *Building Competitiveness in Africa's Agriculture: A Guide to Value Chain Concepts and Applications*, World Bank Publications, Washington, DC, USA, 2009.

[45] R. Kaplinsky and M. Morris, "A handbook for value chain research," 2001.

[46] H. Chesbrough and R. S. Rosenbloom, "The role of the business model in capturing value from innovation: evidence from Xerox Corporation's technology spin-off companies," *Industrial and Corporate Change*, vol. 11, no. 3, pp. 529–555, 2002.

[47] B. K. Ngugi, E. A. Elliot, and C. Blankson, "Toward a conceptual framework explaining the strategic factors responsible for mobile product innovation success in an emerging african market," *Thunderbird International Business Review*, vol. 11, 2020.

[48] R. C. Geibel and M. Manickam, "Comparison of selected startup ecosystems in Germany and in the USA explorative analysis of the startup environments," *GSTF Journal on Business Review (GBR)*, vol. 4, no. 3, 2016.

[49] B. L. Herrmann, J. F. Gauthier, D. Holtschke, R. Berman, and M. Marmer, "The global startup ecosystem ranking 2015," 2015.

[50] D. Isenberg, "The entrepreneurship ecosystem strategy as a new paradigm for economic policy: principles for cultivating entrepreneurship," 2011.

[51] V. K. Vankamamidi, "A conceptual study on startups and T-hubs," *Journal Advances in Business Management*, vol. 3, no. 2, pp. 71–73, 2017.

[52] J. A. C. Baum and B. S. Silverman, "Picking winners or building them? Alliance, intellectual, and human capital as selection criteria in venture financing and performance of biotechnology startups," *Journal of Business Venturing*, vol. 19, no. 3, pp. 411–436, 2004.

[53] T. Hellmann and M. Puri, "Venture capital and the professionalization of start-up firms: empirical evidence," *The Journal of Finance*, vol. 57, no. 1, pp. 169–197, 2002.

[54] J. A. C. Baum, T. Calabrese, and B. S. Silverman, "Don't go it alone: alliance network composition and startups' performance in Canadian biotechnology," *Strategic Management Journal*, vol. 21, no. 3, pp. 267–294, 2000.

[55] G. Walker, B. Kogut, and W. Shan, "Social capital, structural holes and the formation of an industry network," *Organization Science*, vol. 8, no. 2, pp. 109–125, 1997.

[56] Y. Zheng, J. Liu, and G. George, "The dynamic impact of innovative capability and inter-firm network on firm valuation: a longitudinal study of biotechnology start-ups," *Journal of Business Venturing*, vol. 25, no. 6, pp. 593–609, 2010.

[57] C. Liao, "Fuzzy analytical hierarchy process and multi-segment goal programming applied to new product segmented under price strategy," *Computers and Industrial Engineering*, vol. 61, no. 3, pp. 831–841, 2011.

[58] N. J. Rowan and C. M. Galanakis, "Unlocking challenges and opportunities presented by covid-19 pandemic for cross-cutting disruption in agri-food and green deal innovations: quo vadis?" *Science of The Total Environment*, vol. 748, 2020.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2021, Article ID 2529754, 13 pages
https://doi.org/10.1155/2021/2529754

Research Article

Generalized Orthogonal Discrete W Transform and Its Fast Algorithm

Jichao Sun[1,2] and Zhengping Zhang[1]

[1]*College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China*
[2]*College of Computer and Information Engineering, Anhui Polytechnic University, Wuhu 241000, China*

Correspondence should be addressed to Zhengping Zhang; zpzhang@gzu.edu.cn

Received 14 July 2020; Revised 19 November 2020; Accepted 26 November 2020; Published 11 January 2021

Academic Editor: Stefania Tomasiello

Based on the generalized discrete Fourier transform, the generalized orthogonal discrete W transform and its fast algorithm are proposed and derived in this paper. The orthogonal discrete W transform proposed by Zhongde Wang has only four types. However, the generalized orthogonal discrete W transform proposed by us has infinite types and subsumes a family of symmetric transforms. The generalized orthogonal discrete W transform is a real-valued orthogonal transform, and the real-valued orthogonal transform of a real sequence has the advantages of simple operation and facilitated transmission and storage. The generalized orthogonal discrete W transforms provide more basis functions with new frequencies and phases and hence lead to more powerful analysis and processing tools for communication, signal processing, and numerical computing.

1. Introduction

The orthogonal transform is a good mathematical tool and has been used in many applications of digital signal processing, such as harmonic analysis, numerical computation, image processing, data compression, and information hiding [1–6]. Therefore, the method of constructing orthogonal transforms has become an important research topic. The orthogonal transforms include many transforms such as discrete Fourier transform (DFT), discrete Hartley transform, discrete W transform, discrete cosine transform, discrete Walsh transform, and discrete Haar transform.

The DFT is the most widely used and influential transform. Since it was proposed by Fourier in 1822, the Fourier transform has been the most basic analysis method in the analysis of continuous signal and system [7]. In order to adapt to the spectrum analysis by computer, DFT was proposed, which is the approximation of spectrum analysis for continuous time signals. Frequency domain analysis was often superior to time-domain analysis. However, it was impractical to use DFT in the spectrum analysis due to high computational complexity before the advent of the fast algorithm of DFT (FFT). In 1965, Cooley and Tukey [8] published a famous paper on the FFT, which reduced the computation time of DFT by several orders of magnitude. As a result, FFT technology has been widely used in various scientific fields. It broadly promoted the effective combination of engineering and computer technologies. Also, it realized the rapid and effective analysis and processing of engineering problems by enabling use of computers. Fourier transform has a wide range of applications in discrete dynamics, acoustics, optics, and other physics, as well as number theory, combinatorial mathematics, probability, statistics, signal processing, cryptography, and many other fields [9–11].

In order to extend the application scope of DFT, the kernel function of DFT should be generalized to construct as many basis functions with new frequencies and phases as possible.

The kernel function of the DFT is $[\exp(i(2\pi/N))]^{lx}$, where $x = 0, 1, \ldots, N-1$, $l = 0, 1, \ldots, N-1$. If $\omega^N = 1$ and $\omega^n \neq 1$, where $n < N$ (n is a positive integer), ω is called an Nth-order primitive root of unity. It is easy to verify that $\exp(i(2\pi/N))$ is an Nth-order primitive root of unity in the

field of complex numbers ($N \geq 2$). In fact, the base $\exp(i(2\pi/N))$ of the kernel function can be further extended to any Nth-order primitive root of unity in the field of complex numbers, and (l, x) can be extended to $(l+\alpha, x+\beta)$, where α and β can be any real numbers. In this paper, we prove that if ω_N is an Nth-order primitive root of unity in the field of complex numbers ($N \geq 2$), then the row vectors of the transform matrix associated with the kernel function $\omega_N^{(l+\alpha)(x+\beta)}$ are orthogonal to each other, where l and x denote, respectively, the row and column indices of the transform matrix, while the α and β parameters can be any real numbers. The generalized discrete Fourier transform is constructed using the normalized kernel function $(1/\sqrt{N})\omega_N^{(l+\alpha)(x+\beta)}$.

However, since the DFT is a complex-valued transform, a real sequence becomes a complex sequence after DFT. Complex sequences are not as easy to transmit and store as real sequence. Therefore, a real-valued transform was studied to replace the DFT. The discrete Hartley transform (DHT) advanced by Bracewell [12] is a real-valued orthogonal transform. When the DHT is applied to real sequences, it avoids complex operations and speeds up calculations [13]. A transformed real sequence is still a real sequence, which is easy to transmit and store. The DHT and the DFT are related by a simple conversion relationship; that is, the DHT kernel function $\mathrm{cas}((2\pi/N)lx) = \cos((2\pi/N)lx) + \sin((2\pi/N)lx)$, where $x, l = 0, 1, \ldots, N-1$ is the sum of the real and imaginary parts of the DFT kernel function $\exp(i(2\pi/N)lx) = \cos((2\pi/N)lx) + i\sin((2\pi/N)lx)$, where i is the imaginary unit, $x, l = 0, 1, \ldots, N-1$, and $\mathrm{cas}(x)$ denotes $\cos(x) + \sin(x)$ for a real number x. Therefore, the DHT can be used widely instead of DFT [14].

Wang and Hunt [15] developed the discrete W transform (DWT), which is a generalized DHT. There are four types of orthogonal discrete W transform (ODWT), whose kernel function is still the sum of real and imaginary parts of Fourier transform kernel function. DWT and DHT have attracted many scholars' attention due to their advantage compared to DFT in processing real number sequences [16].

DWT and DHT are both discrete Hartley-type transforms. Besides DFT, other complex-valued orthogonal transforms have been recently gaining attention. For each of these complex-valued orthogonal transforms, a real-valued orthogonal transform can be obtained by adding the real and imaginary parts of the kernel function of the complex-valued orthogonal transform together to form the kernel function of the real-valued orthogonal transform. Many of these constructed real-valued orthogonal transforms were found to resemble the DHT and hence are collectively called the discrete Hartley-type transforms [17–20]. These include the Zhang–Hartley transform and the discrete W transform (DWT). Such transforms have been applied in spectral analysis, data compression, convolution, data security, and so on. These transforms have been widely used in communication and signal processing [21–30]. Numerous approaches on fast algorithms for discrete Hartley-type transforms have been reported [31–34].

Thus, for the generalized discrete Fourier transform, its corresponding discrete Hartley-type transform should be considered. Accordingly, the generalized ODWT (GODWT) is proposed in this paper. We show that ODWT is a special case of GODWT. Unlike the ODWT that has only four types, the proposed GODWT has infinite types, subsuming a family of symmetric transforms.

GODWT can replace generalized DFT in wide applications by inheriting the advantages of DHT and DWT. The GODWT is a Hartley-type transform, and it has its fast algorithm. The transformed signal of the real sequence is still of real values, which makes the transmission and storage convenient. Compared with DHT and DWT, GODWT provides a large number of basis functions with new frequency, phase, and a large number of new transformations. It indicates that more transmission methods of information are available when adopting GODWT. In the digital hiding technology, the password can be embedded into transform coefficients. The GODWT is applied to digital hiding technology in the transform domain. The transform space is expanded from ODWT space to GODWT space and the secret key space increases significantly so that the decoding gets more difficult.

GODWT not only provides integer multiples of frequency and phase but also provides various fractional multiples of frequency and phase. The signals in the objective world are various, and it is likely to be a linear combination of several fractional frequency and phase basis functions. If only integer frequency and phase basis function are used for orthogonal decomposition, the number of basis functions will be increased. The proposed fractional multiple frequency and phase basis functions provide a basis for simplifying some problems. The choice of the basis function is actually a subtle matter. For example, if the function $y = x^2$ is expanded by the Maclaurin series in a zero-centered finite interval, only one term arises in the expansion. However, there are an infinite number of terms if the same function is expanded by a trigonometric series. From this example, we can see the importance of the basis function choice for function approximation, spectral analysis, data compression, etc. A problem can be well solved only if it uses basis functions whose frequencies and phases are appropriate for or match the problem. Based on the generalized DFT and GODWT, we have at our disposal a large number of new basis functions with rich frequency and phase information.

In this paper, we use the primitive roots of unity to construct the generalized DFT. Based on the generalized DFT, we also propose and prove the GODWT. Thus, ODWT is expanded from the original four types to infinite types. Moreover, we propose the fast algorithms for computing the generalized DFT and the GODWT. At the end of the paper, the fast algorithm of GODWT and the application of new frequency and phase basis function in communication are illustrated and given as an example. In addition, the key space of GODWT used in digital hiding technology is analyzed. In a word, our generalization provides better mathematical tools for analysis and processing in engineering fields such as communication.

2. Construction of the Generalized DFT Using Primitive Roots of Unity

We show in three stages the construction of the generalized discrete Fourier transform using primitive roots of unity in the complex number field:

(a) If q is a positive integer, q and N are relatively prime; that is, $(q, N) = 1$, and we can know that $\exp(i(2\pi/N)q)$ is the Nth-order primitive root of unity in the complex number field ($N \geq 2$).

In fact, when $n = 1, 2, \ldots, N-1$, n is not divisible by N, that is, $N \nmid n$, and $(q, N) = 1$. Then, the product of q and n is not divisible by N, that is, $N \nmid qn$:

$$\therefore \left[\exp\left(i\frac{2\pi}{N}q\right)\right]^n = \exp\left(i\frac{2\pi}{N}qn\right) \neq 1, \tag{1}$$

when

$$n = N,$$
$$\left[\exp\left(i\frac{2\pi}{N}q\right)\right]^n = \exp(i2\pi q) = 1. \tag{2}$$

Hence, $\exp[i(2\pi/N)q]$ is the Nth-order primitive root of unity in the complex number field.

Obviously, $\exp[-i(2\pi/N)q] = [\exp(i(2\pi/N)q)]^{-1}$ is also an Nth-order primitive root of unity in the complex number field.

(b) If $\omega_N = r \cdot \exp(\mathrm{i}\theta)$ is an Nth-order primitive root of unity, then

$$[r \cdot \exp(i\theta)]^N = 1 = \exp(ik2\pi), \tag{3}$$

where k is any integer:

$$\therefore r = 1,$$
$$\theta = \frac{2k\pi}{N}, \text{ then } \boldsymbol{\omega}_N = \exp\left(i\frac{2k\pi}{N}\right), \tag{4}$$

When $k = 0$, $\omega_N = 1$, and ω_N becomes a first-order primitive root of unity. This root has no practical significance and will not be considered.

When k is a positive integer, we get $\omega_N = \exp(i(k2\pi/N))$, where k and N are relatively prime. Otherwise, if k and N have a common divisor s, $s \neq 1$, then $k = sk_1$, $N = sN_1$, and $N_1 < N$; then

$$(\boldsymbol{\omega}_N)^{N_1} = \left[\exp\left(i\frac{k2\pi}{N}\right)\right]^{N_1} = \left[\exp\left(i\frac{k_1 2\pi}{N_1}\right)\right]^{N_1} = 1, \tag{5}$$

Therefore, ω_N is not an Nth-order primitive root of unity. This contradicts the assumptions.

When k is a negative integer, we similarly get $\omega_N = \exp(i(k2\pi/N))$, where $-k$ and N are relatively prime.

From (a) and (b), we realize that an Nth-order primitive root of unity ω_N in the complex number field can only be $\exp(i(k2\pi/N))$, where k is a positive or negative integer and $|k|$ and N are relatively prime ($N \geq 2$).

(c) If τ is a real number and the conjugate of ω_N^τ is denoted by $\overline{\omega_N^\tau}$, then

$$\overline{\boldsymbol{\omega}_N^\tau} = \overline{\left[\exp\left(i\frac{k2\pi}{N}\right)\right]^\tau} = \overline{\exp\left(i\frac{k2\pi}{N}\tau\right)} = \exp\left(-i\frac{k2\pi}{N}\tau\right) = \left[\exp\left(i\frac{k2\pi}{N}\right)\right]^{-\tau} = \boldsymbol{\omega}_N^{-\tau}. \tag{6}$$

Now, we prove the orthogonality of the rows of the transform matrix that is associated with the kernel function $\omega_N^{(l+\alpha)(x+\beta)}$, where l and x denote the row and column indices of the transform matrix, respectively. The parameters α and β can be any real numbers; ω_N is N^{th}-order primitive root of unity. For any two row vectors $\omega_N^{(l_1+\alpha)(x+\beta)}$ and $\omega_N^{(l_2+\alpha)(x+\beta)}$, where l_1, $l_2 \in \{0, 1, \ldots, N-1\}$, $x = 0, 1, \ldots, N-1$, denote their dot product by $\langle \omega_N^{(l_1+\alpha)(x+\beta)}, \omega_N^{(l_2+\alpha)(x+\beta)} \rangle$:

$$\begin{aligned}
\left\langle \boldsymbol{\omega}_N^{(l_1+\alpha)(x+\beta)}, \boldsymbol{\omega}_N^{(l_2+\alpha)(x+\beta)} \right\rangle &= \sum_{x=0}^{N-1}\left[\boldsymbol{\omega}_N^{(l_1+\alpha)(x+\beta)}\overline{\boldsymbol{\omega}_N^{(l_2+\alpha)(x+\beta)}}\right] \\
&= \sum_{x=0}^{N-1}\left[\boldsymbol{\omega}_N^{(l_1+\alpha)(x+\beta)}\boldsymbol{\omega}_N^{-(l_2+\alpha)(x+\beta)}\right], \\
&= \sum_{x=0}^{N-1}\boldsymbol{\omega}_N^{(l_1-l_2)(x+\beta)}, \\
&= \boldsymbol{\omega}_N^{(l_1-l_2)\beta}\sum_{x=0}^{N-1}\boldsymbol{\omega}_N^{(l_1-l_2)x},
\end{aligned} \tag{7}$$

when $l_1 \neq l_2$, $\omega_N^{(l_1-l_2)} \neq 1$, and $(\omega_N)^N = 1$; therefore,

$$\omega_N^{(l_1-l_2)\beta} \sum_{x=0}^{N-1} \omega_N^{(l_1-l_2)x} = \omega_N^{(l_1-l_2)\beta} \frac{1-\omega_N^{(l_1-l_2)N}}{1-\omega_N^{(l_1-l_2)}} = 0. \quad (8)$$

Thus, the row vectors of the transform matrix associated with the kernel function $\omega_N^{(l+\alpha)(x+\beta)}$ are orthogonal to each other.

When $l_1 = l_2$,

$$\boldsymbol{\omega}_N^{(l_1-l_2)\beta} \sum_{x=0}^{N-1} \boldsymbol{\omega}_N^{(l_1-l_2)x} = N. \quad (9)$$

The modulus of each row vector is $\sqrt{N}$.

If $[U]$ is a complex matrix of order N and $[U][U]^T = [U]^T[U] = [E]$, [U] is called unitary matrix, where $[U]^T$ is the transposed conjugate matrix of $[U]$, and $[E]$ is a unit matrix. Therefore, the $N \times N$ matrix generated by the kernel function $(1/\sqrt{N})\,\omega_N^{(l+\alpha)(x+\beta)} = (1/\sqrt{N})\exp[\pm i(2\pi/N)q(l+\alpha)(x+\beta)]$ is a unitary matrix, where l, $x = 0, 1, \ldots, N-1$, and each of the parameters α and β can be any real number. The kernel function $(1/\sqrt{N})\omega_N^{(l+\alpha)(x+\beta)}$ lies at the l-th row and the x-th column of the unitary matrix.

When

$$\alpha = -(2l + \alpha'),$$

$$\frac{1}{\sqrt{N}}\exp\left[i\frac{2\pi}{N}q(l+\alpha)(x+\beta)\right] = \frac{1}{\sqrt{N}}\exp\left[-i\frac{2\pi}{N}q(l+\alpha')(x+\beta)\right]. \quad (10)$$

Therefore, only the kernel function $(1/\sqrt{N})\exp[i(2\pi/N)q(l+\alpha)(x+\beta)]$ constructed by the primitive root of unity $\omega_N = \exp(i(2\pi/N)q)$ should be considered.

The transform with the kernel function $1/\sqrt{N}\exp[i(2\pi/N)q(l+\alpha)(x+\beta)]$ is expressed as

$$F(l) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x)\boldsymbol{\omega}_N^{(l+\alpha)(x+\beta)} = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x)\exp\left[i\frac{2\pi}{N}q(l+\alpha)(x+\beta)\right], \text{where,} \quad l = 0, 1, \ldots N-1. \quad (11)$$

The transform in (11) is called a generalized discrete Fourier transform, and its inverse transform is

$$f(x) = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} F(l)\left[\omega_N^{-1}\right]^{(l+\alpha)(x+\beta)}, \text{where,} \quad x = 0, 1, \ldots N-1. \quad (12)$$

3. Construction of the GODWT

The DWT with two parameters, α and β, is defined as follows:

$$F(l) = \sqrt{\frac{2}{N}} \sum_{x=0}^{N-1} f(x)\sin\left[\frac{\pi}{4} + \frac{2\pi}{N}(l+\alpha)(x+\beta)\right] = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x)\operatorname{cas}\left[\frac{2\pi}{N}(l+\alpha)(x+\beta)\right], \quad (13)$$

where $x = 0, 1, \ldots, N-1$, $l = 0, 1, \ldots, N-1$, and DWT is orthogonal for four (α, β) values, that is, $(\alpha, \beta) = \{(0, 0), ((1/2), 0), (0, (1/2)), ((1/2), (1/2))\}$. The orthogonal discrete W transform has only the four types. We denote the four transforms associated with these parameter value pairs by DWT-I, DWT-II, DWT-III, and DWT-IV, respectively, and we call them ODWT.

The parameter q can be added to the expression in (13) to obtain

$$F(l) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x)\operatorname{cas}\left[\frac{2\pi}{N}q(l+\alpha)(x+\beta)\right], \quad \text{where, } (x = 0, 1, \ldots, N-1), (l = 0, 1, \ldots, N-1). \quad (14)$$

At this point, many meaningful orthogonal transforms can be constructed. Two lemmas are proved as follows.

Lemma 1. *If p is positive integer $(p, N) = 1$, $\beta = (h/2p)$, and both h and m are integers, then*

$$\sum_{x=0}^{N-1} \sin\frac{2\pi}{N}pm(x+\beta) = 0. \quad (15)$$

Proof. If $N \nmid m$, $(p, N) = 1$, $\therefore N \nmid pm$, $\exp(i(2\pi/N)pm) \neq 1$,

$$\therefore 1-\exp\left(i\frac{2\pi}{N}pm\right)\neq 0, \tag{16}$$

From (16) and (17), we get $\sum_{x=0}^{N-1}\exp(i(2\pi/N)pmx)=0$. And,

$$\left[1-\exp\left(i\frac{2\pi}{N}pm\right)\right]\sum_{x=0}^{N-1}\exp\left(i\frac{2\pi}{N}pmx\right)=1-\exp(i2\pi pm)=0. \tag{17}$$

$$\therefore \sum_{x=0}^{N-1}\exp\left[i\frac{2\pi}{N}pm(x+\beta)\right]=\exp\left(i\frac{2\pi}{N}pm\beta\right)\cdot\sum_{x=0}^{N-1}\exp\left(i\frac{2\pi}{N}pmx\right)=0, \therefore \sum_{x=0}^{N-1}\sin\frac{2\pi}{N}pm(x+\beta)=0. \tag{18}$$

If m is divisible by N, that is, $N|m$, $m=N\mathrm{d}$,

$$\sum_{x=0}^{N-1}\sin\frac{2\pi}{N}pm(x+\beta)=\sum_{x=0}^{N-1}\sin\left(\frac{\pi}{N}\right)2pN\,d\left(x+\frac{h}{2p}\right)=\sum_{x=0}^{N-1}\sin\pi d(2px+h)=0. \tag{19}$$

Q.E.D. □

Lemma 2. *A real or complex sequence $f(x)$, where $x=0,1,\ldots N-1$, can be regarded as a function defined on the set $J_N=\{0,1,\ldots,N-1\}$. A finite Abelian group can be formed through modulo-N addition on J_N. This group is called a residual-class additive group and is denoted by Z_N.*

Next, we derive below the conditions for constructing a real-valued orthogonal system by adding the real and imaginary parts of a complex-valued orthogonal system together.

Let the complex-valued function set on Z_N, that is,

$$\{\Phi_k(x)=R_k(x)+iN_k(x)|k=0,1,\ldots,N-1\}, \tag{20}$$

be a normalized orthogonal function system, where $R_k(x)$ and $N_k(x)$ are the real and imaginary parts of $\Phi_k(x)$, respectively. Then, the condition for the real-valued function set,

$$\{\Psi_k(x)=R_k(x)+N_k(x)|k=0,1,\ldots,N-1\}, \tag{21}$$

to be a normalized orthogonal function system is

$$\sum_{x=0}^{N-1}R_{k_1}(x)N_{k_2}(x)=0, \tag{22}$$

where k_1, $k_2\in\{0,1,\ldots,N-1\}$.

Proof. $\{\Phi_k(x)\}$ is a normalized orthogonal system. Hence, for any k_1, $k_2\in\{0,1,\ldots,N-1\}$,

$$\mathrm{Im}\left[\sum_{x=0}^{N-1}\Phi_{k_1}(x)\overline{\Phi_{k_2}(x)}\right]=0, \tag{23}$$

which can be rewritten as

$$\sum_{x=0}^{N-1}\left(-R_{k_1}(x)N_{k_2}(x)+R_{k_2}(x)N_{k_1}(x)\right)=0. \tag{24}$$

Hence, we find that

$$\sum_{x=0}^{N-1}R_{k_1}(x)N_{k_2}(x)=\sum_{n=0}^{N-1}R_{k_2}(x)N_{k_1}(x), \tag{25}$$

$$\sum_{x=0}^{N-1}\Psi_{k_1}(x)\Psi_{k_2}(x)=\sum_{x=0}^{N-1}\Phi_{k_1}(x)\overline{\Phi_{k_2}(x)}+\sum_{x=0}^{N-1}\left(R_{k_1}(x)N_{k_2}(x)+R_{k_2}(x)N_{k_1}(x)\right)+i\sum_{x=0}^{N-1}\left(R_{k_1}(x)N_{k_2}(x)-R_{k_2}(x)N_{k_1}(x)\right). \tag{26}$$

From (25) and (26), we can get

$$\sum_{x=0}^{N-1}\Psi_{k_1}(x)\Psi_{k_2}(x)=\sum_{x=0}^{N-1}\Phi_{k_1}(x)\overline{\Phi_{k_2}(x)}+2\sum_{x=0}^{N-1}R_{k_1}(x)N_{k_2}(x). \tag{27}$$

From (27), we find that, for any k_1, $k_2\in\{0,1,\ldots,N-\}$, the condition $\sum_{x=0}^{N-1}R_{k_1}(x)N_{k_2}(x)=0$ is the necessary and sufficient condition for $\{\Psi_k(x)=R_k(x)+N_k(x),k=0,1,\ldots,N-1\}$ to be a real normalized orthogonal function system. □

3.1. The Generalized Orthogonal Discrete W Transform. The unitary matrix with the kernel function $(1/\sqrt{N})\exp[i(2\pi/N)q(l+\alpha)(x+\beta)]$ is constructed as discussed above. The N row vectors of an $N\times N$ unitary matrix form a normalized orthogonal system. The sum of the real and imaginary parts of the kernel function $(1/\sqrt{N})\exp[i(2\pi/N)q(l+\alpha)(x+\beta)]$ is given as

$$\frac{1}{\sqrt{N}}\left\{\cos\left[\frac{2\pi}{N}q(l+\alpha)(x+\beta)\right]+\sin\left[\frac{2\pi}{N}q(l+\alpha)(x+\beta)\right]\right\}=\frac{1}{\sqrt{N}}\operatorname{cas}\left[\frac{2\pi}{N}q(l+\alpha)(x+\beta)\right], \tag{28}$$

and this sum is used as the kernel function of a new real matrix. According to Lemma 2, we know that the necessary and sufficient conditions for the orthonormality of the N row vectors of the real matrix can be formulated as

$$\frac{1}{N}\sum_{x=0}^{N-1}\cos\left[\frac{2\pi}{N}q(l_1+\alpha)(x+\beta)\right]\sin\left[\frac{2\pi}{N}q(l_2+\alpha)(x+\beta)\right]=0,\quad \text{for any } l_1,l_2\in\{0,1,\ldots,N-1\}, \tag{29}$$

Let

$$\sigma(\alpha,\beta)=\frac{1}{N}\sum_{x=0}^{N-1}\cos\left[\frac{2\pi}{N}q(l_1+\alpha)(x+\beta)\right]\sin\left[\frac{2\pi}{N}q(l_2+\alpha)(x+\beta)\right]. \tag{30}$$

Using the product-to-sum formula (i.e., the Prosthaphaeresis formula), from (30), we can get

$$\sigma(\alpha,\beta)=\frac{1}{2N}\sum_{x=0}^{N-1}\left\{\sin\left[\frac{2\pi}{N}q(l_1+l_2+2\alpha)(x+\beta)\right]-\sin\left[\frac{2\pi}{N}q(l_1-l_2)(x+\beta)\right]\right\}. \tag{31}$$

If $q=q_1q_2$, both q_1 and q_2 are positive integers, $\because(q,N)=1$, $\because(q_2,N)=1$; let $\alpha=(g/2q_1)$ and $\beta=(h/2q_2)$, where g and h can be any integers; by substitution in (31), we get

$$\sigma\left(\frac{g}{2q_1},\frac{h}{2q_2}\right)=\frac{1}{2N}\sum_{x=0}^{N-1}\left\{\sin\left[\frac{2\pi}{N}q_2q_1\left(l_1+l_2+2\frac{g}{2q_1}\right)\left(x+\frac{h}{2q_2}\right)\right]-\sin\left[\frac{2\pi}{N}q_2q_1(l_1-l_2)\left(x+\frac{h}{2q_2}\right)\right]\right\}, \tag{32}$$

$\therefore q_1(l_1+l_2+2\alpha)=q_1(l_1+l_2+2(g/2q_1))=q_1(l_1+l_2)+g$ is an integer and $q_1(l_1-l_2)$ is also an integer.

According to Lemma 1, we get

$$\sigma\left(\frac{g}{2q_1},\frac{h}{2q_2}\right)=\frac{1}{2N}\sum_{x=0}^{N-1}\left\{\sin\left[\frac{2\pi}{N}q_2q_1\left(l_1+l_2+2\frac{g}{2q_1}\right)\left(x+\frac{h}{2q_2}\right)\right]-\sin\left[\frac{2\pi}{N}q_2q_1(l_1-l_2)\left(x+\frac{h}{2q_2}\right)\right]\right\}=0, \tag{33}$$

Then, when $\alpha = (g/2q_1)$ and $\beta = (h/2q_2)$, equation (29) holds, and the N functions,

$$\frac{1}{\sqrt{N}}\operatorname{cas}\left[\frac{2\pi}{N}q(l+\alpha)(x+\beta)\right] = \frac{1}{\sqrt{N}}\operatorname{cas}\left[\frac{2\pi}{N}q\left(l+\frac{g}{2q_1}\right)\left(x+\frac{h}{2q_2}\right)\right], \quad \text{where } (l = 0, 1, \ldots, N-1), \tag{34}$$

form a normalized orthogonal basis, where the independent variable $x = 0, 1, \ldots, N-1$.

The real-valued orthogonal matrix is obtained with the kernel functions in (34), where l and x are the row and column indices of the matrix. The transform

$$F(l) = \frac{1}{\sqrt{N}}\sum_{x=0}^{N-1} f(x)\operatorname{cas}\left[\frac{2\pi}{N}q\left(l+\frac{g}{2q_1}\right)\left(x+\frac{h}{2q_2}\right)\right], \quad \text{where } (l = 0, 1, \ldots N-1), \tag{35}$$

is GODWT, and its inverse transform is

$$f(x) = \frac{1}{\sqrt{N}}\sum_{l=0}^{N-1} F(l)\operatorname{cas}\left[\frac{2\pi}{N}q\left(l+\frac{g}{2q_1}\right)\left(x+\frac{h}{2q_2}\right)\right], \quad \text{where } (l = 0, 1, \ldots, N-1). \tag{36}$$

If s_1 is the factor of $2\,q_1$ and s_2 is the factor of $2\,q_2$, we get $2q_1 = t_1 s_1$ and $2q_2 = t_2 s_2$. So, parameter α of (28) and (29) becomes $\alpha = (g/2q_1) = (g/t_1 s_1)$. Similarly, parameter β of (28) and (29) becomes $\beta = (h/2q_2) = (h/t_2 s_2)$. Then, we set $g = t_1 g_1$ and $h = t_2 h_2$, where g_1 and h_2 can be any integers. Finally, we obtain $\alpha = (g_1/s_1)$ and $\beta = (h_2/s_2)$. Thus, the N row vectors of the real matrix with the kernel functions $(1/\sqrt{N})\operatorname{cas}[(2\pi/N)q(l + (g_1/s_1))(x + (h_2/s_2))]$ form a normalized orthogonal basis. If you change g_1 and h_2 into g and h, or any other two letters, the new letters express the same settings as long as they are any integers.

Let $\alpha = (g/s_1)$ and $\beta = (h/s_2)$, where g and h can be any integers; if $q = q_1 q_2 = 1$, the factors s_1 of $2q_1$ and s_2 of $2q_2$ can only be taken as 1 or 2. When $(\alpha, \beta) \in \{(0, 0), ((1/2), 0), (0(1/2)), ((1/2)(1/2))\}$, the four transforms are given by the kernel function $(1/\sqrt{N})\operatorname{cas}[(2\pi/N)\mathrm{q}(l+\alpha)(x+\beta)]$, and it is obvious that they are DWT-I, DWT-II, DWT-III, and DWT-IV, respectively, that is, ODWT.

For the GODWT, when $q = 1$, it reduces to the ODWT, and when $q \neq 1$, a new family of orthogonal transforms arises. This family greatly expands the ODWT and also subsumes a family of symmetric transforms.

If q contains a factor s^2, we can construct a transform matrix such that $\alpha = (f/2s)$ and $\beta = (f/2s)$, where f can be any integer. This transform matrix includes two special cases, namely, $(\alpha = (f/2), \beta = (f/2))$ and $(\alpha = (f/s), \beta = (f/s))$. In these cases, the real-valued orthogonal matrix is a symmetric matrix, and the forward and inverse transforms have the same form.

4. Fast Algorithms for the Generalized Discrete Fourier and GODWT

4.1. The Fast Algorithm for the Generalized Discrete Fourier Transform. Let $f(x)$ be a sequence, where $x = 0, 1, \ldots, N-1$, and let ω_N be an Nth-order primitive root of unity in the complex number field. Then, the transform of $f(x)$ by the kernel function $(1/\sqrt{N})\omega_N^{(l+\alpha)(x+\beta)}$ (the generalized DFT) is expressed as

$$F(l) = \frac{1}{\sqrt{N}}\sum_{x=0}^{N-1} f(x)\omega_N^{(l+\alpha)(x+\beta)} = \left\{\frac{1}{\sqrt{N}}\sum_{x=0}^{N-1}\left[f(x)\omega_N^{\alpha x}\right]\omega_N^{lx}\right\}\omega_N^{(l+\alpha)\beta}, \quad \text{where } (l = 0, 1, \ldots, N-1). \tag{37}$$

From (37), we can identify two steps for the transform of $f(x)$ with the kernel function $(1/\sqrt{N})\omega_N^{(l+\alpha)(x+\beta)}$:

(1) When N is a composite number (nonprime, $N \geq 2$), $f(x)\omega_N^{\alpha x}$ is transformed by the kernel function

$(1/\sqrt{N})\omega_N^{lx}$. This is adopted for the generalized fast Fourier transform (FFT) [35].

(2) The result from the first step is multiplied by $\omega_N^{(l+\alpha)\beta}$.

So, when N is a composite number, there is a fast algorithm for computing the transform of $f(x)$ with the kernel function $(1/\sqrt{N})\omega_N^{(l+\alpha)(x+\beta)}$.

If ω_N is an Nth-order primitive root of unity in the complex field, then ω_N^{-1} is also an Nth-order primitive root of unity in the complex field. Similarly, when N is a composite number, a fast algorithm exists for computing the inverse transform

$$f(x) = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} F(l)\left[\boldsymbol{\omega}_N^{-1}\right]^{(l+\alpha)(x+\beta)} = \frac{1}{\sqrt{N}} \sum_{l=0}^{N-1} F(l)\overline{\omega}_N^{(l+\alpha)(x+\beta)}, \quad \text{where } (l = 0, 1, \ldots, N-1). \tag{38}$$

4.2. The Fast Algorithm for the GODWT. Let $[\Phi]$ be a unitary matrix: $[\Phi] = [R] + i[N]$, where $[R]$ and $[N]$ are real matrices and $[\Psi] = [R] + [N]$ forms an orthogonal matrix. Then, the inverse transform matrix of $[\Phi]$ is its conjugate transposed matrix $\overline{[\Phi]^T} = [R]^T - i[N]^T$. The inverse transform matrix of $[\Psi]$ is its transposed matrix $[\Psi]^T = [R]^T + [N]^T$.

If $[L]$ is a real-valued column vector, then the actions of the matrices on this vector are as follows:

$$[\Phi][L] = ([R] + i[N])[L] = [R][L] + i[N][L], \tag{39}$$

$$[\Psi][L] = ([R] + [N])[L] = [R][L] + [N][L], \tag{40}$$

$$\overline{[\Phi]^T}[L] = \left([R]^T - i[N]\right)[L] = [R]^T[L] - i[N]^T[L], \tag{41}$$

$$[\Psi]^T[L] = \left([R]^T + [N]^T\right)[L] = [R]^T[L] + [N]^T[L], \tag{42}$$

From (39) and (40), we see that $[\Psi][L]$ is the sum of the real and imaginary parts of $[\Phi][L]$. Moreover, from (41) and (42), we see that $[\Psi]^T[L]$ is the real part minus the imaginary part of $\overline{[\Phi]^T}[L]$.

For forward and inverse transforms of the GODWT, the transformed sequences are all real value sequences. Using the above relations of $[\Psi][L]$ and $[\Phi][L]$ as well as $[\Psi]^T[L]$ and $\overline{[\Phi]^T}[L]$, we can obtain fast algorithms for computing the GODWT and its inverse transform. However, the advantages of real-valued operations can be brought into play only by developing direct fast algorithms of real-valued orthogonal transforms on real sequences. This is one of the problems in the anticipated future work on the GODWT.

Example 1. Let $N = 8$, $q = 9$, $q_1 = 3$, $q_2 = 3$, $g = 1$, and $h = 1$. Therefore, $\alpha = (g/2q_1) = (1/6)$, and $\beta = (h/2q_2) = (1/6)$. The transform is

$$F(l) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x)\operatorname{cas}\left[\frac{2\pi}{N} q\left(l + \frac{g}{2q_1}\right)\left(x + \frac{h}{2q_2}\right)\right] = \frac{\sqrt{2}}{4} \sum_{x=0}^{7} f(x)\operatorname{cas}\left[\frac{2\pi \times 9}{8}\left(l + \frac{1}{6}\right)\left(x + \frac{1}{6}\right)\right], \quad \text{where } f(x) \text{ is the real sequence.} \tag{43}$$

The fast algorithm of the transform in (43) is given as follows.

Due to the fact that the kernel function $(\sqrt{2}/4)\operatorname{cas}[((2\pi \times 9)/8)(l + (1/6))(x + (1/6))]$ is the sum of the real and imaginary parts of the generalized DFT kernel function $(\sqrt{2}/4)\omega^{(l+(1/6))(x+(1/6))}$, where $\omega = \exp[i((2\pi \times 9)/8)]$, we firstly calculate the transform of $f(x)$ with the kernel function $(\sqrt{2}/4)\omega^{(l+(1/6))(x+(1/6))}$:

$$F''(l) = \frac{\sqrt{2}}{4} \sum_{x=0}^{7} f(x)\omega^{(l+(1/6))(x+(1/6))} = \left\{\sum_{x=0}^{7}[f(x)\omega^{(1/6)x}]\omega^{lx}\right\} \frac{\sqrt{2}}{4}\omega^{(l+(1/6))(1/6)}. \tag{44}$$

In order to get $F''(l)$, first, we calculate $f'(x)$, $f'(x) = f(x)\omega^{(1/6)x}$, where $x = 0, 1, \ldots, 7$. Second, $f'(x)$ is transformed with the kernel function ω^{lx} into $F'(l)$; then we calculate $F''(l)$, $F''(l) = (\sqrt{2}/4)F'(l)\omega^{(l+(1/6))(1/6)}$, where $l = 0, 1, \ldots, 7$. We get $F(l)$ after the real and the imaginary parts of $F''(l)$ are added.

The matrix $[\Phi]$ is formed with the kernel function ω^{lx}, $x = 0, 1, \ldots, 7$, $l = 0, 1, \ldots, 7$. It can be decomposed into the

product of a sparse matrix according to the method in [35], so it has a fast algorithm $[\Phi] = [F_1][F_2][F_3]$,

$$[\Phi] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & \omega & \omega^2 & \omega^3 & \omega^4 & \omega^5 & \omega^6 & \omega^7 \\ 1 & \omega^2 & \omega^4 & \omega^6 & 1 & \omega^2 & \omega^4 & \omega^6 \\ 1 & \omega^3 & \omega^6 & \omega & \omega^4 & \omega^7 & \omega^2 & \omega^5 \\ 1 & \omega^4 & 1 & \omega^4 & 1 & \omega^4 & 1 & \omega^4 \\ 1 & \omega^5 & \omega^2 & \omega^7 & \omega^4 & \omega & \omega^6 & \omega^3 \\ 1 & \omega^6 & \omega^4 & \omega^2 & 1 & \omega^6 & \omega^4 & \omega^2 \\ 1 & \omega^7 & \omega^6 & \omega^5 & \omega^4 & \omega^3 & \omega^2 & \omega \end{bmatrix},$$

$$[F_1] = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \omega & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \omega^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \omega^3 \\ 1 & \omega^4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \omega^5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \omega^6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & \omega^7 \end{bmatrix}, \tag{45}$$

where $\omega^4 = -1$.

$$[F_2] = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \omega^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \omega^2 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \omega^6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & \omega^6 \end{bmatrix},$$

$$[F_3] = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \end{bmatrix}. \tag{46}$$

By using the sparse matrix, it is easy to draw a signal flow diagram that a column vector is transformed by the matrix $[\Phi]$. The FFT flow diagram is depicted in Figure 1.

We calculate the inverse transform of $F(l) = (\sqrt{2}/4) \sum_{x=0}^{7} f(x)\mathrm{cas}[(2\pi \times 9/8)(l + (1/6))(x + (1/6))]$. First, we calculate the inverse transform of the transform associated with the kernel $(\sqrt{2}/4)\omega^{(l+(1/6))(x+(1/6))}$ as follows:

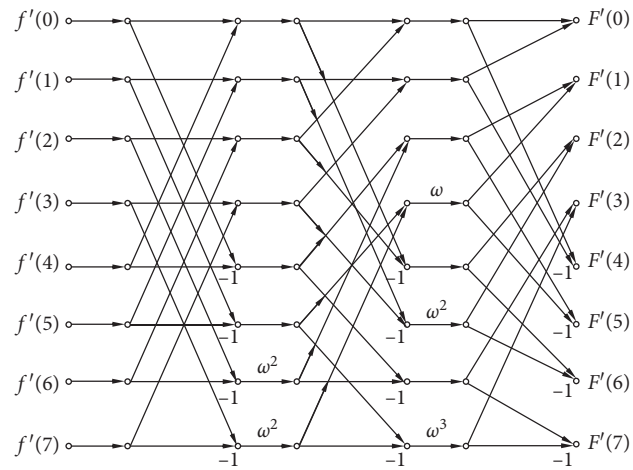


Figure 1: The FFT flow diagram.

$$\begin{aligned} f_2(x) &= \frac{\sqrt{2}}{4} \sum_{l=0}^{7} F(l)\omega^{-(l+(1/6))(x+(1/6))} \\ &= \left\{ \sum_{l=0}^{7} [F(l)\overline{\omega}^{(1/6)l}]\overline{\omega}^{lx} \right\} \frac{\sqrt{2}}{4}\overline{\omega}^{(x+(1/6))(1/6)}, \end{aligned} \tag{47}$$

where $\overline{\omega}$ is the conjugate of ω.

In order to get $f_2(x)$, first, $F_1(l)$ is computed as $F_1(l) = F(l)\overline{\omega}^{(1/6)l}$. Then, the FFT of $F_1(l)$ is computed using a kernel function $\overline{\omega}^{lx}$. The transform matrix $[\Phi_1]$ formed by kernel function $\overline{\omega}^{lx}$ is the conjugate transpose matrix $[\Phi]^T$ of $[\Phi]$. $[\Phi]$ is the symmetric matrix. So $[\Phi_1] = \overline{[\Phi]}$. In the product of sparse matrix of $[\Phi]$, that is, $[F_1] * [F_2] * [F_3]$ and its corresponding signal flow diagram, ω is replaced with $\overline{\omega}$, the product of the sparse matrix decomposed by $[\Phi_1]$ and its corresponding signal flow diagram can be obtained. The input and output of the signal flow diagram are $F_1(l)$ and $f_1(x)$, respectively. Then, given $f_2(x) = (\sqrt{2}/4) f_1(x)\omega^{(x+(1/6))(1/6)}$, $f(x)$ is obtained using the real part minus the imaginary part of $f_2(x)$.

Another method of computing $\sum_{l=0}^{7}[[F(l)\overline{\omega}^{(1/6)l}]\overline{\omega}^{lx}$ is as follows:

$\sum_{l=0}^{7}[[F(l)\overline{\omega}^{(1/6)l}]\overline{\omega}^{lx} = \sum_{l=0}^{7}[[\overline{F(l)}\omega^{(1/6)l}]\omega^{lx}$, so, in order to calculate the FFT of $F(l)\overline{\omega}^{(1/6)l}$ with kernel function $\overline{\omega}^{lx}$, we can firstly calculate the FFT of $\overline{F(l)}\omega^{(1/6)l}$ with kernel function ω^{lx} and then take the conjugate. The product of sparse matrix and signal flow graph of Φ can be directly used for FFT with kernel function ω^{lx}.

In this example, the transform $F(l) = (\sqrt{2}/4)\sum_{x=0}^{7} f(x)\mathrm{cas}[((2\pi \times 9)/8)(l + (1/6))(x + (1/6))$ is symmetric transform; thus, the inverse transform is the same as the positive transform. The inverse transform is

$$f(x) = \frac{\sqrt{2}}{4} \sum_{l=0}^{7} F(l)\mathrm{cas}\left[\frac{2\pi \times 9}{8}\left(l + \frac{1}{6}\right)\left(x + \frac{1}{6}\right)\right]. \tag{48}$$

Therefore, we can use the same fast algorithm as the positive transform.

4.3. Examples of the GODWT Application

(1) In communication field, the communication between different users using different frequency bands is frequency division multiplexing, and the communication between different users relying on different address codes is code division multiplexing. The code division multiplexing and frequency division multiplexing can be combined by using GODWT. We still take the primitive unit root $\omega = \exp[i(2\pi \times 9/8)]$, and three kernel functions of GODWT, that is, $(\sqrt{2}/4)\mathrm{cas}[((2\pi \times 9)/8)xl]$, $(\sqrt{2}/4)$ cas $[((2\pi \times 9)/8)(x + (1/3))l]$, and $(\sqrt{2}/4)\mathrm{cas}[((2\pi \times 9)/8)(x + (2/3))l]$ can be obtained. For the three transmitted code sequences, that is, $a(x)$, $b(x)$, and $c(x)$, $x = 0, 1, \ldots, 7$, the transforms of them are calculated, respectively, to get $A(l)$, $B(l)$, and $C(l)$, where $l = 0, 1, \ldots, 7$.

$$A(l) = \frac{\sqrt{2}}{4}\sum_{x=0}^{7} a(x)\mathrm{cas}\left[\frac{2\pi \times 9}{8}xl\right],$$
$$B(l) = \frac{\sqrt{2}}{4}\sum_{x=0}^{7} b(x)\mathrm{cas}\left[\frac{2\pi \times 9}{8}\left(x + \frac{1}{3}\right)l\right], \tag{49}$$
$$C(l) = \frac{\sqrt{2}}{4}\sum_{x=0}^{7} c(x)\mathrm{cas}\left[\frac{2\pi \times 9}{8}\left(x + \frac{2}{3}\right)l\right],$$

$A(l) = (\sqrt{2}/4)\sum_{x=0}^{7} a(x)\mathrm{cas}[((2\pi \times 9)/8)xl]$; the equation is written in matrix forms as

$$[A(0), A(1), \ldots, A(7)] = \frac{\sqrt{2}}{4}[a(0), a(1), \ldots, a(7)]\begin{bmatrix} \mathrm{cas}\left[\frac{9\pi}{4}0 \times 0\right] & \mathrm{cas}\left[\frac{9\pi}{4}0 \times 1\right] & \ldots & \mathrm{cas}\left[\frac{9\pi}{4}0 \times 7\right] \\ \mathrm{cas}\left[\frac{9\pi}{4}1 \times 0\right] & \mathrm{cas}\left[\frac{9\pi}{4}1 \times 1\right] & \ldots & \mathrm{cas}\left[\frac{9\pi}{4}1 \times 7\right] \\ \vdots & \vdots & \vdots & \vdots \\ \mathrm{cas}\left[\frac{9\pi}{4}7 \times 0\right] & \mathrm{cas}\left[\frac{9\pi}{4}7 \times 1\right] & \ldots & \mathrm{cas}\left[\frac{9\pi}{4}7 \times 7\right] \end{bmatrix}. \tag{50}$$

Let the carrier of code division multiplexing be $Z_{10}, Z_{11}, \ldots, Z_{17}$:

$$\begin{aligned} Z_{10} &= \left[\mathrm{cas}\left[\frac{9\pi}{4}0 \times 0\right] \ \mathrm{cas}\left[\frac{9\pi}{4}0 \times 1\right] \ \ldots \ \mathrm{cas}\left[\frac{9\pi}{4}0 \times 7\right]\right], \quad \text{the angular frequency of } Z_{10} \text{ is } \frac{9\pi}{4} \times 0, \\ Z_{11} &= \left[\mathrm{cas}\left[\frac{9\pi}{4}1 \times 0\right] \ \mathrm{cas}\left[\frac{9\pi}{4}1 \times 1\right] \ \ldots \ \mathrm{cas}\left[\frac{9\pi}{4}1 \times 7\right]\right], \quad \text{the angular frequency of } Z_{11} \text{ is } \frac{9\pi}{4} \times 1, \\ &\vdots \quad \vdots \quad \ldots \quad \vdots \quad \ldots \quad \ldots \quad \vdots \\ Z_{17} &= \left[\mathrm{cas}\left[\frac{9\pi}{4}7 \times 0\right] \ \mathrm{cas}\left[\frac{9\pi}{4}7 \times 1\right] \ \ldots \ \mathrm{cas}\left[\frac{9\pi}{4}7 \times 7\right]\right], \quad \text{the angular frequency of } Z_{10} \text{ is } \frac{9\pi}{4} \times 7 \end{aligned} \tag{51}$$

The carriers $Z_{10}Z_{11}, \ldots, Z_{17}$ are periodic and repetitive. The amounts of information carried by each carrier wave over a period are constant, which are $(\sqrt{2}/4)a(0)$, $(\sqrt{2}/4)a(1)\ldots(\sqrt{2}/4)a(7)$, respectively.

The matrix equation above can be written as

$$[A(0), A(1), \ldots, A(7)] = \frac{\sqrt{2}}{4}\{a(0)Z_{10} + a(1)Z_{11} + \cdots + a(7)Z_{17}\}. \tag{52}$$

In order to transmit $A(l) = (\sqrt{2}/4) \sum_{x=0}^{7} a(x)\text{cas}[((2\pi \times 9)/8)xl]$, $l = 0, 1, \ldots, 7$, the angular frequency of carriers is from $(9\pi/4) \times 0$ to $(9\pi/4) \times 7$. It is called code division multiplexing where we have transmitting $A(l)$, $l = 0, 1, \ldots, 7$ instead of transmitting $a(x)$, $x = 0, 1, \ldots, 7$. Both the input and output are eight real numbers. In contrast, if we use the DFT, the output is eight complex numbers. The transmission of eight complex numbers requires the transmission of sixteen real numbers. Accordingly, the transmission efficiency of GODWT is two times that of DFT. Transmitting $B(l)$ and $C(l)$ instead of transmitting $b(x)$ and $c(x)$ is also code division multiplexing, but they all have different carriers. $B(l) = (\sqrt{2}/4) \sum_{x=0}^{7} a(x)\text{cas}[((2\pi \times 9)/8)(x + (1/3))l]$, $l = 0, 1, \ldots, 7$; when transmitting $B(l)$, the angular frequency of carriers is from $(9\pi/4) \times (0 + (1/3))$ to $(9\pi/4) \times (7 + (1/3))$. $C(l) = (\sqrt{2}/4) \sum_{x=0}^{7} c(x)\text{cas}[((2\pi \times 9)/8)(x + (2/3))l]$, $l = 0, 1, \ldots, 7$; when transmitting $C(l)$, the angular frequency of carriers is from $(9\pi/4) \times (0 + (2/3))$ to $(9\pi/4) \times (7 + (2/3))$.

When they are transmitted together, it is frequency division multiplexing. The receiver would receive their sum:

$$A(l) + B(l) + C(l) = \frac{\sqrt{2}}{4} \sum_{x=0}^{7} a(x)\text{cas}\left[\frac{2\pi \times 9}{8} xl\right] + \frac{\sqrt{2}}{4} \sum_{x=0}^{7} b(x)\text{cas}\left[\frac{2\pi \times 9}{8}\left(x + \frac{1}{3}\right)l\right] + \frac{\sqrt{2}}{4} \sum_{x=0}^{7} c(x)\text{cas}\left[\frac{2\pi \times 9}{8}\left(x + \frac{2}{3}\right)l\right]. \tag{53}$$

These three components are separated with a filter, and then $a(x)b(x)$,$c(x)$ were restored using the inverse GODWT.

Time-division multiplexing (TDM) is also a method to transmit multiway signals. Time-division multiplexing is used primarily for digital signals but may be applied in analog multiplexing in which two or more signals or bit streams are transferred appearing simultaneously as subchannels in one communication channel but are physically taking turns on the channel. Each signal appears on the line only a fraction of time in an alternating pattern. The time domain is divided into several recurrent time slots of fixed length.

Transmitting $A(l)$, $B(l)$, and $C(l)$ with different time slots, they can be separated at the receiving end, and then $a(x)$, $b(x)$, and $c(x)$ were restored using the inverse GODWT. This realizes the combination of code-division multiplexing and time-division multiplexing. It shows that the proposed GODWT provides more means for the transmission in communication, compared with DHT and ODWT.

(2) In the digital hiding technology, the password can be embedded into transform coefficients.

The secret key space of ODWT is (N, α, β), where $(\alpha, \beta) = \{(0, 0), ((1/2), 0), (0, (1/2)), ((1/2), (1/2))\}$, and N is a positive integer. Actually, N cannot be big; if N is bigger, it is not easy to calculate. The secret key space of GODWT is (N, q, α, β). It includes the secret key space of ODWT. Because the value of q is not limited in principle and the parameter value pairs (α, β) changed from 4 pairs of ODWT to infinity pairs of GODWT: $(\alpha, \beta) = ((g/2q_1), (h/2q_2))$ where $q_1 q_2 = q$, g and h can be any integers, the secret key space of GODWT has been greatly generalized.

5. Conclusions

The kernel function of the generalized DFT is $(1/\sqrt{N})\exp[i(2\pi/N)q(l + \alpha)(x + \beta)]$, where $(q, N) = 1$, and parameters α and β can be any real numbers. When $l = 0, 1, \ldots, N$-1, these N functions form a complex-valued normalized orthogonal system.

For the kernel function $(1/\sqrt{N})\exp[i(2\pi/N)q(l + \alpha)(x + \beta)]$ of the generalized DFT, the real and imaginary parts are added together leading to the functions:

$(1/\sqrt{N})$ cas $[(2\pi/N)q(l + \alpha)(x + \beta)] = (1/\sqrt{N})\{\cos[(2\pi/N)q(l + \alpha)(x + \beta)] + \sin[(2\pi/N)q(l + \alpha)(x + \beta)]\}$, $(l = 0, 1, \ldots, N$-1). If $q = q_1 q_2$, both q_1 and q_2 are positive integers; let $\alpha = (g/2q_1)$, $\beta = (h/2q_2)$, where g and h are any integers. These N functions form a real normalized orthogonal system, where the independent variable $x = 0, 1, \ldots, N$-1. The transform with the kernel function $(1/\sqrt{N})$ cas $[(2\pi/N)q(l + (g/2q_1))(x + (h/2q_2))]$ is GODWT. For setting the values of α and β, if the factor μ_1 of $2q_1$ replaces $2q_1$ and the factor μ_2 of $2q_2$ replaces $2q_2$, the conclusion is still valid.

When N is a composite number, based on the fast algorithms of the generalized DFT and its inverse transform, we get the fast algorithm of the GODWT and its inverse transform.

GODWT can replace the generalized DFT in wide applications. The transformed signal of the real sequence is still of real values, which makes the transmission and storage convenient. Also, GODWT provides a large number of basis functions with a new frequency, phase, and a large number of new transforms and subsumes a family of symmetric transforms; their forward and reverse transformations can be implemented by the same computer programs or hardware.

It can be seen that GODWT can provide more transmission means for communication, better security for digital hiding technology, and many new methods for data

compression. Hence, it can be used as a more powerful analysis and processing tool for communication, signal processing, and numerical computing.

Abbreviations

N, n, q, q_1, q_2, p, s, t:	Positive integers
d, f, g, g_1, h, h_2, k, m:	Integers
α, β, τ:	Real numbers
ω:	Complex numbers
$[\Phi]$, $[\Psi]$, $[R]$, $[N]$, $[F_1]$, $[F_2]$, $[F_3]$:	The matrices
$[L]$:	Real-valued column vector
ω_N:	Nth-order primitive root of unity in the complex number field
$\overline{\omega}$:	Conjugate of ω
$\overline{[\Phi]}$:	Conjugate matrix of the matrix $[\Phi]$
$[\Phi]^T$:	Transposed matrix of the matrix $[\Phi]$
$\overline{[\Phi]^T}$:	Conjugate transposed matrix of the matrix $[\Phi]$.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the projects of the Chinese Natural Science Foundation (Grant no. 61865002), Anhui Province Higher Education Promotion Plan (Grant no. TSKJ2017B30), the Natural Science Fund of Anhui Province (Grant no. 1908085ME128), the Natural Science Foundation of Education Department of Anhui Province (Grant no. KJ2018A0123), and Shanghai Key Laboratory of Materials Laser Processing and Modification (Grant no. MLPM2017-3).

References

[1] H. Andrews, "Orthogonal transforms for digital signal processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 5, 438 pages, 1976.

[2] D. Puchala, "Involutory parametric orthogonal transforms of cosine-walsh type with application to data encryption," *Advances in Intelligent Systems and Computing*, vol. 689, 2017.

[3] Y. Önal and Ü. Ç. Turhal, "The orthogonal Hilbert-Huang transform application in voltage flicker analysis," in *Proceedings of the 4th International Conference on Power Engineering, Energy and Electrical Drives*, pp. 700–704, Istanbul, Turkey, May 2013.

[4] A. Zemliachenko, V. Lukin, and M. Alhihi, "Analysis of approaches to hyperspectral data compression based on orthogonal transforms," in *Proceedings of the 14th International Conference on Advanced Trends in Radioelecrtronics, Telecommunications and Computer Engineering (TCSET)*, pp. 909–914, Slavske, Ukraine, February 2018.

[5] R. Ezhilarasi and K. Venkatalakshmi, "Low complexity orthogonal transforms for low cost image/video codec design," *Journal of Computational and Theoretical Nanoscience*, vol. 15, no. 3, pp. 859–865, 2018.

[6] L. Cheng, Z. Jiang, and Z. Zhang, "The generalized unified computation of multidimensional discrete orthogonal transforms," *Science in China Series: Information Sciences*, vol. 44, pp. 401–411, 2001.

[7] A. G. Kjelaas, "The fourier transform: properties and applications," *Surveillance of Environmental Pollution and Resources by Electromagnetic Waves*, vol. 45, pp. 329–351, 1978.

[8] J. W. Cooley and J. W. Tukey, "An algorithm for the machine computation of complex fourier series," *Mathematics of Computation*, vol. 19, pp. 297–301, 1965.

[9] R. Brenner, "Numerical computation of the response of piezoelectric composites using Fourier transform," *Physical Review B Condensed Matter*, vol. 79, no. 18, Article ID 184106, 2009.

[10] R. Bracewell, "The Fourier transform and its applications," *American Journal of Physics*, vol. 34, 2002.

[11] W. Zhou, X. Luo, B. Chen, Y. Zhang, and X. Cai, "Estimation of particle depth from two defocused images using the Fourier transform," *Particuology*, vol. 49, pp. 48–54, 2020.

[12] R. N. Bracewell, "Discrete Hartley transform," *Journal of the Optical Society of America*, vol. 73, no. 12, pp. 1832–1835, 1983.

[13] H. S. Hou, "The fast Hartley transform algorithm," *IEEE Transactions on Computers*, vol. 36, no. 2, pp. 147–156, 1987.

[14] X. Ouyang, J. Jin, G. Jin, and Z. Wang, "Low complexity discrete Hartley transform precoded OFDM for peak power reduction," *Electronics Letters*, vol. 48, no. 2, pp. 90-91, 2012.

[15] Z. Wang and B. R. Hunt, "The discreteW transform," *Applied Mathematics and Computation*, vol. 16, no. 1, pp. 19–48, 1985.

[16] N. C. Hu, H. I. Chang, and O. K. Ersoy, "Generalized discrete Hartley transforms," *IEEE Transactions on Signal Processing*, vol. 40, no. 12, pp. 2931–2940, 1992.

[17] A. Bortoletti and C. D. Fiore, "On a set of matrix algebras related to discrete Hartley-type transforms," *Linear Algebra and Its Applications*, vol. 366, pp. 65–85, 2003.

[18] S. Pei, C. Tseng, M. Yeh et al., "Discrete fractional Hartley and fourier transforms," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, vol. 45, no. 6, pp. 665–675, 1998.

[19] C. Moraga, "Analysis of Mosaics by means of the Chrestenson and Zhang-Hartley transforms," in *Proceedings The Nineteenth International Symposium on Multiple-Valued Logic*, pp. 421–427, Guangzhou, China, May 1989.

[20] Z. Wang, "Fast algorithms for the discrete W transform and for the discrete Fourier transform," *IEEE Transactions on Acoustics Speech & Signal Processing*, vol. 32, no. 4, pp. 803–816, 2003.

[21] E. Chilton, M. Barbarosou, and I. Paraskevas, "Hartley transform and the use of the Whitened Hartley spectrum as a tool for phase spectral processing," *The Journal of Engineering*, vol. 2015, no. 3, pp. 95–101, 2015.

[22] J. K. Mandal and S. K. Ghosal, "Separable discrete Hartley transform based invisible watermarking for color image authentication (SDHTIWCIA)," *Advances in Computing and Information Technology*, vol. 177, pp. 767–776, 2013.

[23] R. S. Sunder, C. Eswaran, and N. Sriraam, "Medical image compression using 3-D Hartley transform," *Computers in Biology and Medicine*, vol. 36, pp. 958–973, 2006.

[24] N. X. Thao and H. T. Van Anh, "On the Hartley-fourier sine generalized convolution," *Mathematical Methods in the Applied Sciences*, vol. 37, no. 15, pp. 2308–2319, 2014.
[25] M. Al-Gharabally and A. F. Almutairi, "Frequency-domain subcarrier diversity receiver for discrete Hartley transform OFDM systems," *EURASIP Journal on Wireless Communications and Networking*, vol. 2019, 2019.
[26] M. R. Abuturab, "An asymmetric single-channel color image encryption based on Hartley transform and gyrator transform," *Optics and Lasers in Engineering*, vol. 69, pp. 49–57, 2015.
[27] Z. Liu, Y. Zhang, W. Liu, F. Meng, Q. Wu, and S. Liu, "Optical color image hiding scheme based on chaotic mapping and Hartley transform," *Optics and Lasers in Engineering*, vol. 51, no. 8, pp. 967–972, 2013.
[28] W. A. Martins and P. S. R. Diniz, "Memoryless block transceivers with minimum redundancy based on Hartley transforms," *Signal Processing*, vol. 91, no. 2, pp. 240–251, 2011.
[29] X. Ouyang, J. Jin, G. Jin, and Z. Wang, "Discrete Hartley transform based SFBC-OFDM transceiver design with low complexity," in *Proceedings of the 2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 2744–2749, Shanghai, China, April 2013.
[30] P. Rajavel, "Directional Hartley transform and content based image retrieval," *Signal Processing*, vol. 90, no. 4, pp. 1267–1278, 2010.
[31] I. Prots' ko, "Algorithm of efficient computation of generalised discrete Hartley transform based on cyclic convolutions," *IET Signal Processing*, vol. 8, pp. 301–308, 2013.
[32] S. Bouguezel, M. O. Ahmad, and M. N. S. Swamy, "New parametric discrete fourier and Hartley transforms and algorithms for fast computation," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, pp. 562–575, 2010.
[33] G. Bi, "On computation of the discrete W transform," *IEEE Transactions on Signal Processing*, vol. 47, pp. 1450–1453, 1999.
[34] V. Kober, "Fast algorithms for the computation of sliding discrete Hartley transforms," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2937–2944, 2007.
[35] J. A. Glassman, "A generalization of the fast fourier transform," *IEEE Transactions on Computers*, vol. 19, no. 2, pp. 105–116, 1970.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2020, Article ID 8848031, 9 pages
https://doi.org/10.1155/2020/8848031

Research Article

A Novel Regret Theory-Based Decision-Making Method Combined with the Intuitionistic Fuzzy Canberra Distance

Haiping Ren,[1] **Yunxiao Gao**,[2] **and Tonghua Yang**[3]

[1]*Teaching Department of Basic Subjects, Jiangxi University of Science and Technology, Nanchang 330013, China*
[2]*School of Science, Jiangxi University of Science and Technology, Ganzhou 341000, China*
[3]*School of Vocational Education and Technology, Jiangxi Agricultural University, Nanchang 330045, China*

Correspondence should be addressed to Haiping Ren; chinarhp@163.com

Received 17 August 2020; Revised 10 September 2020; Accepted 24 September 2020; Published 22 October 2020

Academic Editor: Stefania Tomasiello

In practical decision-making, the behavior factors of decision makers often affect the final decision-making results. Regret theory is an important behavioral decision theory. Based on the regret theory, a novel decision-making method is proposed for the multiattribute decision-making problem with incomplete attribute weight information, and the attribute values are expressed by Atanassov intuitionistic fuzzy numbers. At first, a new distance of intuitionistic fuzzy sets is put forward based on the traditional Canberra distance. Then, we utilize it for the definition of the regret value (rejoice) for the attribute value of each alternative with the corresponding values of the positive point (negative point). The objective of this method is to maximize the comprehensive perceived utility of the alternative set by the decision maker. The optimal attribute weight vector is solved, and the optimal comprehensive perceived utility value of each alternative is obtained. Finally, according to the optimal comprehensive perceived utility value, the rank order of all alternatives is concluded.

1. Introduction

Since Professor Zadeh introduced the concept of fuzzy set in 1965, it has been successfully applied in many fields such as intelligent control, military engineering, economic prediction, and decision-making [1–5]. Zadeh's fuzzy sets have been proved to be an effective tool to deal with fuzzy and imprecise problems [6, 7]. However, in the process of solving some decision-making problems, due to the limitation of the time, energy, or incomplete knowledge of decision makers, decision makers often hesitate, which makes the evaluation results show three aspects: affirmation, negation, and hesitation. Traditional fuzzy sets cannot describe such problems very well, so Professor Atanassov [8] extended fuzzy sets in 1986 and introduced the concept of intuitionistic fuzzy set (IFS). By introducing the parameter of nonmembership degree, IFS can express the information of affirmation and negation at the same time; furthermore, it can describe the fuzzy concept of "not this or that," and then it can describe the hesitation and uncertainty of the decision maker's judgment [9, 10]. Because of this, it can depict the fuzzy essence of the objective world more delicately than Zadeh's fuzzy set in the processing mode. In recent years, the research on the IFS theory has attracted great attention of scholars and has been applied to the fields of economic decision-making [11, 12], medical diagnosis [13, 14], image processing [15, 16], pattern recognition [17, 18], fault tree analysis [19, 20], and so on. Some other extensions of Zadeh's fuzzy set such as picture fuzzy set [6] and rough set [21] all have received great attention and have many successful applications in practice.

Most of the existing intuitionistic fuzzy MADM methods are based on the expected utility theory, which assumes that decision makers are completely rational. However, decision makers often have subjective preferences, such as psychological and behavioral factors when making decision. So, it is important to consider the subjective preferences of decision makers in the decision process. As an important behavioral decision-making theory, regret theory was firstly proposed by Bell [22] and Loomes and Sugden [23]. In recent years,

the research and application of the regret theory have attracted many scholars' attention [24–27]. Regret theory holds that decision makers are concerned about the possible results if they choose other schemes while considering the results of schemes. If they find that they can get better results by choosing other schemes, they will regret them psychologically. Otherwise, they will be happy. Therefore, when making a decision, the decision maker will estimate the regret or rejoice that the decision may produce in advance and try to avoid choosing the plan of expected regret; that is to say, the decision maker is regret-averse. In this way, the perceived utility value of the decision maker includes two parts: the utility value of the current result and the perceived utility value after comparing with other possible results, "regret-rejoice" value. Chorus [28] and Qu et al. [29] pointed out that the regret theory has some advantages over the cumulative prospect theory in application. For example, in decision-making, reference points need not be given, and few parameters in the calculation formula were involved in decision-making, which makes the calculation simpler [30]. In the decision-making model, regret theory replaces the expected utility theory, and it is in line with the objective reality of human beings [31].

Many information measures are proposed for fuzzy sets, such as entropy, similarity, and distance [32]. In the classical regret theory, the deviation involving two numbers can be directly measured in accordance with the absolute value of the difference between two numbers. To measure the difference between two IFSs, we are required to define the distance between two IFSs. Although many intuitionistic fuzzy distance measures have been constructed, some existing distance measures have counter-intuitionistic special cases, so it is very important to develop new improved intuitionistic fuzzy distance measures. Canberra distance, as a classical distance measure, has been widely applied in image processing, pattern recognition, and other fields based on the exact number. Taking into account this distance measure, this paper develops a novel intuitionistic fuzzy distance based on the Canberra distance and further applies it to develop a new intuitionistic fuzzy decision-making method combined with the regret theory.

The structure of this paper is as follows: Section 2 first introduces the concept of IFS and then provides some preliminaries of the regret theory. Section 3 puts forward a new intuitionistic fuzzy distance measure based on the traditional Canberra distance. Section 4 develops the intuitionistic fuzzy multiattribute decision-making (MADM) method based on the regret theory combined with the proposed intuitionistic fuzzy Canberra distance. Section 5 provides an example, which explains the new method through the example analysis. Finally, Section 6 is the conclusion of this paper.

2. Preliminaries

Some basic concepts and properties of IFSs and regret theory are reviewed in this section.

Definition 1 (see [8]). Let $X = \{x_1, x_2, \ldots, x_n\}$ be a universal set. A set U is called an IFS in X if $U = \{< x_i, \mu_U(x_i), \nu_U(x_i) > \mid x_i \in X\}$. Here, $\mu_U(x_i)$ and $\nu_U(x_i)$ are the membership degree and nonmembership degree of x_i, respectively. They satisfy $\mu_U(x_i)$, $\nu_U(x_i)$, $\mu_U(x_i) + \nu_U(x_i) \in [0, 1]$ for $\forall x_i \in X$. Let $\pi_U(x_i) = 1 - \mu_U(x_i) - \nu_U(x_i)$; then, $\pi_U(x_i)$ is called the hesitancy degree of x_i. Sometimes, IFS $U = \{< x_i, \mu_U(x_i), \nu_U(x_i) > \mid x_i \in X\}$ is also expressed as $U = \{< x_i, \mu_U(x_i), \nu_U(x_i), \pi_U(x_i) > \mid x_i \in X\}$.

Remark 1. Grzegorzewski [33] introduced the concept of intuitionistic fuzzy number (IFN) as an extension of the IFS in the continuous case, and intuitionistic fuzzy numbers have many applications in the engineering field [34]. In order to avoid confusion and for convenience, if there is only one element in X, we call U an Atanassov intuitionistic fuzzy number (AIFN). Each AIFN has a physical interpretation, for example, if A=<0.6, 0.2, 0.2>, then $\mu_A = 0.6$, $\nu_A = 0.2$, and $\pi_A = 0.2$, which can be interpreted as "the vote for resolution is 6 in favor, 2 against, and 2 abstention."

In actual decision-making process, most of the decision makers are not so rational, so the decision maker's behavior factors need to be considered when making a decision. The prospect theory and regret theory are put forward in this context. In the increasingly complex, modern, political, and economic environment, decision makers need to consider not only the results obtained after choosing a certain scheme but also the possible decision results after assuming that other alternatives are chosen. In the regret theory, the perceived utility function is composed of two parts: the utility function of current decision-making results and the regret-rejoice function compared with other decision-making results. Let a and b, respectively, represent the results that can be obtained by selecting scheme A and scheme B. Then, the perceived utility of decision makers for scheme A is

$$u(a, b) = v(a) + R(v(a) - v(b)). \tag{1}$$

Among them, $v(\theta)$ represents the utility value of scheme θ and $R(v(a) - v(b))$is called regret-rejoice value. If $R(v(a) - v(b))$ is positive, then it is called a rejoice value, which indicates the extent to which the decision maker is glad to choose the scheme or give up the scheme. If $R(v(a) - v(b))$ is negative, then it is called a regret value, which indicates the extent to which the decision maker regrets to choose the scheme or give up the scheme. Obviously, the regret gratification function $R(\cdot)$ should be monotonically increasing and concave, i.e., it satisfies $R'(\cdot) > 0$, $R''(\cdot) < 0$, and $R(0) = 0$. Loomes and Sugden [23] pointed out that regret-rejoice function $R(\cdot)$can be expressed as follows:

$$R(\Delta v) = 1 - \exp(-\delta \Delta v). \tag{2}$$

Here, $\delta > 0$ is the regret avoidance coefficient of the decision maker, and the greater δ related to, the larger the regret avoidance degree of the decision maker. Δv is the difference between the utility value of any two schemes.

Figure 1 shows the image of the regret-rejoice function with different values.

Let $A_i\,(i = 1, 2, \ldots, m)$ be i-th alternatives, and a_i is the result of alternative A_i. According to the regret theory, in decision analysis, when the positive ideal point is taken as the reference point, the decision-making evaluation value will not be greater than the positive ideal point, and at this time, the decision maker will regret. When the negative ideal point is taken as the reference point, the decision-making evaluation value will not be less than the negative ideal point, and at this time, the decision maker is happy. Note that x_{ij} is the attribute evaluation value of scheme A_i under the evaluation attribute o_j given by the decision maker; then, according to Loomes and Sugden [23], the regret value of each attribute evaluation value x_{ij} of scheme A_i is related to the corresponding attribute valuex_j^+ of the positive ideal point, and the joy value of each attribute evaluation value x_{ij} of scheme A_i is related to the corresponding attribute value x_j^- of the negative ideal point. The gratification values can be expressed as

$$\begin{aligned} R_{ij}^1 &= 1 - \exp\left(\delta\left|x_{ij} - x_j^+\right|\right), \\ R_{ij}^2 &= 1 - \exp\left(-\delta\left|x_{ij} - x_j^-\right|\right), \end{aligned} \tag{3}$$

where $\delta > 0$ is the regret avoidance coefficient of decision makers.

A large number of psychological studies have shown that regret, as a negative emotion, has a stronger effect on utility than rejoice. Therefore, the decision maker's comprehensive regret-rejoice value for the evaluation value of scheme A_i under the evaluation attribute o_j is

$$R_{ij} = R_{ij}^1 + R_{ij}^2 = 2 - \exp\left(\delta\left|x_{ij} - x_j^+\right|\right) - \exp\left(-\delta\left|x_{ij} - x_j^-\right|\right). \tag{4}$$

According to Loomes and Sugden [23], power function $v(x) = x^\alpha$, $0 < \alpha < 1$, is used as a utility function of the attribute value in this paper. The greater the degree of risk aversion of decision makers is, the smaller α is. α is called the risk aversion coefficient of decision makers. It can be proved that if $\Delta v_0 > 0$, there is $|R(-\delta\Delta v_0)| > R(\Delta v_0)$. This shows that compared with Δv_0, decision maker's psychological perception is more sensitive to $-\Delta v_0$, that is, decision makers are regret-averse.

3. A New Distance Based on the Canberra Distance

In this section, we will propose a new distance measure between two IFSs based on the Canberra distance. The

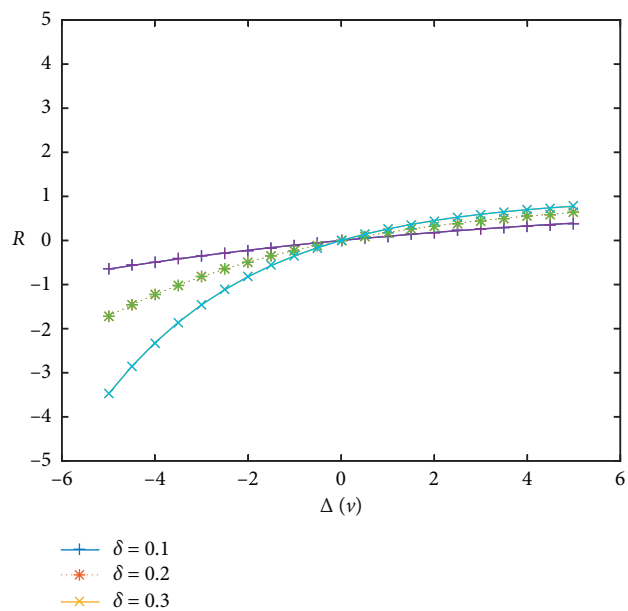


Figure 1: Regret-rejoice function, $R\Delta v$.

Canberra distance of two real vectors $x = (x_1, x_2, \ldots, x_n)$ and $y = (y_1, y_2, \ldots, y_n)$is defined as follows (Perlibakas [35]):

$$d(x, y) = \sum_{i=1}^{n} \frac{|x_i - y_i|}{|x_i| + |y_i|}. \tag{5}$$

As an important information measure, Canberra distance has been successfully applied in image processing, medicine, and other fields [36–38]. Due to the fact that the denominator is zero, the numerical value is meaningless. Then, we propose a revised version of the Canberra distance measure as follows:

$$\sum_{i=1}^{n} d(x, y) = \sum_{i=1}^{n} \frac{|x_i - y_i|}{2 + |x_i| + |y_i|}. \tag{6}$$

Note that constant 2 can be changed as any other positive numbers.

Let $X = \{x_1, x_2, \ldots, x_n\}$ be a universal set. Then, for two given IFSs, $A = \{< x_i, \mu_A(x_i), \nu_A(x_i), \pi_A(x_i) > \mid x_i \in X\}$and $B = \{< x_i, \mu_B(x_i), \nu_B(x_i), \pi_B(x_i) > \mid x_i \in X\}$, and the new intuitionistic fuzzy information measure based on Canberra distance $d(A, B)$ is constructed as follows:

$$d(A, B) = \frac{1}{n} \sum_{i=1}^{n} \left[\frac{|\mu_A(x_i) - \mu_B(x_i)|}{2 + \mu_A(x_i) - \mu_B(x_i)} + \frac{|\nu_A(x_i) + \nu_B(x_i)|}{2 + \nu_A(x_i) + \nu_B(x_i)} + \frac{|\pi_A(x_i) + \pi_B(x_i)|}{2 + \pi_A(x_i) + \pi_B(x_i)} \right]. \tag{7}$$

Next, we will prove $d(A,B)$ is a valid distance measure. In this section, let R^* be a set of nonnegative real numbers.

Lemma 1. *Let $a,b,c \in R^*$ and $a \le b \le c$. Then,*

$$\begin{aligned} &\text{(i)} \quad \frac{|a-c|}{2+a+c} \ge \frac{|a-b|}{2+a+b}, \\ &\text{(ii)} \quad \frac{|a-c|}{2+a+c} \ge \frac{|b-c|}{2+b+c}, \\ &\text{(iii)} \quad \frac{|a-c|}{2+a+c} \le \frac{|a-b|}{2+a+b} + \frac{|b-c|}{2+b+c}. \end{aligned} \tag{8}$$

Lemma 2. *Let $a,b,c \in R^*$ and $d_1(a,b) = (|a-b|)/(2+a+b)$; then,*

$$d_1(a,b) \le d_1(b,c) + d_1(a,c). \tag{9}$$

Proof. For the case $0 \le a \le b \le c$, according to Lemma 1, we have

(i) $d_1(a,c) \ge d_1(a,b)$; then, $d_1(a,b) \le d_1(a,c) + d_1(b,c)$

(ii) $d_1(a,c) \ge d_1(b,c)$; then, $d_1(b,c) \le d_1(a,c) + d_1(a,b)$

(iii) $d_1(a,c) \le d_1(a,b) + d_1(b,c)$

That is, $d_1(\cdot,\cdot)$ satisfies trigonometric inequality. And we can easily prove that $d_1(\cdot,\cdot)$satisfies trigonometric inequality in other cases using a similar reasoning process. Then, the conclusion is proved. □

Theorem 1. *Let $A = \{<x_i, \mu_A(x_i), \nu_A(x_i)> \mid x_i \in X\}$ and $B = \{<x_i, \mu_B(x_i), \nu_B(x_i)> \mid x_i \in X\}$ be two IFSs in $X = (x_1, x_2, \ldots, x_n)$. Then, $d(A,B)$ defined in (7) is a valid distance measure between A and B. That is, $d(A,B)$ satisfies the following properties:*

(i) $d(A,B) \ge 0$

(ii) $d(A,B) = d(B,A)$

(iii) $d(A,B) \le d(B,C) + d(A,C)$, *for any IFSs $A, B,$ and C*

Proof

(i) Obviously, $d(A,B) \ge 0$.

(ii) $d(B,A) = (1/n)\sum_{i=1}^{n}[(|\mu_B(x_i) - \mu_A(x_i)|/2 + \mu_B(x_i) + \mu_A(x_i)) + (|\nu_B(x_i) - \nu_A(x_i)|/2 + \nu_B(x_i) + \nu_A(x_i)) + (|\pi_B(x_i) - \pi_A(x_i)|/2 + \pi_B(x_i) + \pi_A(x_i))] = (1/n)\sum_{i=1}^{n}[(|\mu_A(x_i) - \mu_B(x_i)|/2 + \mu_A(x_i) + \mu_B(x_i)) + (|\nu_A(x_i) - \nu_B(x_i)|/2 + \nu_A(x_i) + \nu_B(x_i)) + (|\pi_A(x_i) - \pi_B(x_i)|/2 + \pi_A(x_i) + \pi_B(x_i))] = d(A,B)$.

(iii) If A, B, and C are three IFSs, $A = \{<x_i, \mu_A(x_i), \nu_A(x_i)> \mid x_i \in X\}$, $B = \{<x_i, \mu_B(x_i), \nu_B(x_i)> \mid x_i \in X\}$, and $C = \{<x_i, \mu_C(x_i), \nu_C(x_i)> \mid x_i \in X\}$, then by (i) of Lemma 2, we have

$$\begin{aligned} d_1(\mu_A(x_i) + \mu_B(x_i)) &\le d_1(\mu_B(x_i) + \mu_C(x_i)) + d_1(\mu_A(x_i) + \mu_C(x_i)), \\ d_1(\nu_A(x_i) + \nu_B(x_i)) &\le d_1(\nu_B(x_i) + \nu_C(x_i)) + d_1(\nu_A(x_i) + \nu_C(x_i)), \\ d_1(\pi_A(x_i) + \pi_B(x_i)) &\le d_1(\pi_B(x_i) + \pi_C(x_i)) + d_1(\pi_A(x_i) + \pi_C(x_i)), \end{aligned} \tag{10}$$

while

$$d(A,B) = \frac{1}{n}\sum_{i=1}^{n}[d_1(\mu_A(x_i) + \mu_B(x_i)) + d_1(\nu_A(x_i) + \nu_B(x_i)) + d_1(\pi_A(x_i) + \pi_B(x_i))]. \tag{11}$$

Consequently, $d(A,B) \le d(B,C) + d(A,C)$

Then, we complete the proof of Theorem 1.

If we consider the important degree of x_i $(i = 1, 2, \ldots, n)$ and let w_i $(i = 1, 2, \ldots, n)$ be the important degree of x_i, $(i = 1, 2, \ldots, n)$, which satisfies $w_i \in [0,1]$ and $\sum_{i=1}^{n} w_i = 1$, then we can get a weighted distance $d_W(A,B)$ between A and B as follows:

$$d_W(A,B) = \sum_{i=1}^{n} w_i \left[\frac{|\mu_A(x_i) - \mu_B(x_i)|}{2 + \mu_A(x_i) - \mu_B(x_i)} + \frac{|\nu_A(x_i) + \nu_B(x_i)|}{2 + \nu_A(x_i) + \nu_B(x_i)} + \frac{|\pi_A(x_i) + \pi_B(x_i)|}{2 + \pi_A(x_i) + \pi_B(x_i)}\right]. \tag{12}$$

□

Remark 2. If $w_i = 1/n$, $(i = 1, 2, \ldots, n)$, then $d_W(A, B) = d(A, B)$. Obviously, $d_W(A, B)$ is also a valid distance, and the proof process is similar to $d(A, B)$ in Theorem 1.

4. A New Regret Theory-Based Decision-Making Method

In this section, we will put forward a new intuitionistic fuzzy MADM method based on the regret theory combined with the above proposed distance. The detail decision process is shown in Figure 2.

For an intuitionistic fuzzy MADM problem, for the convenience of description, the following symbols represent the set or quantity in the decision:

$X = \{x_1, x_2, \ldots, x_n\}$: the set of m alternatives $(m \geq 2)$.

$O = \{o_1, o_2, \ldots, o_n\}$: the set of n attributes $(n \geq 2)$, where o_j represents the jth attribute.

$\mathbf{w} = (w_1, w_2, \ldots, w_n)^T$: the vector of attributes' weights. This is because in the decision-making process, different attributes usually have different importance. Here, w_j is the weight information of attribute o_j $(j = 1, 2, \ldots, n)$, satisfying $w_j \geq 0$ $(j = 1, 2, \ldots, n)$ and $\sum_{j=1}^{n} w_j = 1$. When the attribute weight information is partially known, the set of mathematical expressions that record the known partial weight information is denoted by H.

$x_{ij} = <\mu_{ij}, \nu_{ij}>$: the evaluation value of alternative x_i given by the decision maker under attribute o_j. The numbers μ_{ij} and ν_{ij} show the degree of satisfaction and dissatisfaction of the decision maker with the value of the alternative x_i under the index o_j, respectively. They satisfy $0 \leq \mu_{ij} \leq 1$, $0 \leq \nu_{ij} \leq 1$, and $0 \leq \mu_{ij} + \nu_{ij} \leq 1$.

Now, we can get an intuitionistic fuzzy decision-making matrix $\widetilde{X} = (x_{ij})_{m \times n}$. It is required to determine the order of alternatives and choose the optimal alternative.

Now, we propose a new decision-making method based on the regret theory. The decision maker's comprehensive regret-rejoice value of the evaluation value $x_{ij} = <\mu_{ij}, \nu_{ij}>$ of scheme x_i under the evaluation attribute o_j is

$$R_{ij} = R_{ij}^1 + R_{ij}^2 = 2 - \exp\left[\delta d\left(x_{ij}, x_j^+\right)\right] - \exp\left[-\delta d\left(x_{ij}, x_j^-\right)\right]. \tag{13}$$

In this paper, function $v_{ij}(x_{ij}) = S(x_{ij})^\alpha$, $0 < \alpha < 1$, is used as a utility function of the attribute value, and $S(x)$ is the score function of the AIFN. Then, the decision maker's perception utility function of the corresponding attribute value x_{ij} of scheme x_i can be expressed as

$$\begin{aligned} F_{ij} = v_{ij} + R_{ij} &= 2 + S\left(x_{ij}\right)^\alpha - \exp\left[\delta d\left(x_{ij}, x_j^+\right)\right] \\ &\quad - \exp\left[-\delta d\left(x_{ij}, x_j^-\right)\right]. \end{aligned} \tag{14}$$

Next, we discuss the method to determine the attribute weight of intuitionistic fuzzy MADM. Let H be the set of known weight information. For each scheme x_i, its comprehensive perceived utility function is

$$\begin{aligned} F(x_i) &= \sum_{j=1}^{n} w_j F_{ij} \\ &= \sum_{j=1}^{n} w_j \left\{2 + S\left(x_{ij}\right)^\alpha - \exp\left[\delta d\left(x_{ij}, x_j^+\right)\right]\right. \\ &\quad \left. - \exp\left[-\delta d\left(x_{ij}, x_j^-\right)\right]\right\}. \end{aligned} \tag{15}$$

The weight should be determined so that the greater the comprehensive perceived utility, the better the scheme x_i is. Therefore, the following optimization model can be established, and its objective function is

$$\max F = \left(F(x_1), F(x_2), \ldots, F(x_m)\right). \tag{16}$$

According to the fact that "the greater the comprehensive perceived utility, the better the scheme" and the fair competition among the schemes aimed at the maximization of the comprehensive perceived utility of the decision maker to the scheme set, the optimization model for solving the attribute weight is established as follows:

$$\begin{aligned} \max V &= \sum_{i=1}^{m} F(x_i) \\ &= \sum_{i=1}^{m} \sum_{j=1}^{n} w_j \left\{2 + S\left(x_{ij}\right)^\alpha - \exp\left[\delta d\left(x_{ij}, x_j^+\right)\right]\right. \\ &\quad \left. - \exp\left[-\delta d\left(x_{ij}, x_j^-\right)\right]\right\}, \\ \text{s.t.} &\quad \begin{cases} \mathbf{w} \in \mathbf{H}, \\ \sum_{j=1}^{n} w_j = 1, \\ w_j \geq 0, \quad j = 1, 2, \ldots, n. \end{cases} \end{aligned} \tag{17}$$

The optimal weight vector $\mathbf{w}^* = (w_1^*, w_2^*, \ldots, w_m^*)^T$ can be obtained by solving the above model with MATLAB or LINGO software.

Thus, it can be seen that the optimal comprehensive perceived utility value of the decision maker to scheme x_i is

$$\begin{aligned} H_i &= \sum_{j=1}^{n} w_j^* \left\{2 + S\left(x_{ij}\right)^\alpha - \exp\left[\delta d\left(x_{ij}, x_j^+\right)\right]\right. \\ &\quad \left. - \exp\left[-\delta d\left(x_{ij}, x_j^-\right)\right]\right\}. \end{aligned} \tag{18}$$

Finally, according to the comparison of the optimal comprehensive perceived utility value, the ranking results of all schemes can be obtained. The larger the value H_i is, the better the corresponding alternative x_i is.

Next, the calculation steps of the MADM method based on the regret theory are given as follows:

Step 1: calculate the score $S_{ij} = \mu_{ij} - \nu_{ij}$ of attribute value $x_{ij} = <u_{ij}, \nu_{ij}>$, and get the score matrix $S = (S_{ij})_{m \times n}$.

Step 2: determine the positive and negative ideal point.

The positive ideal point is defined as $x^+ = (x_1^+, x_2^+, \ldots, x_n^+)$, where $x_j^+ = <1, 0>$, $j = 1, 2, \ldots, n$

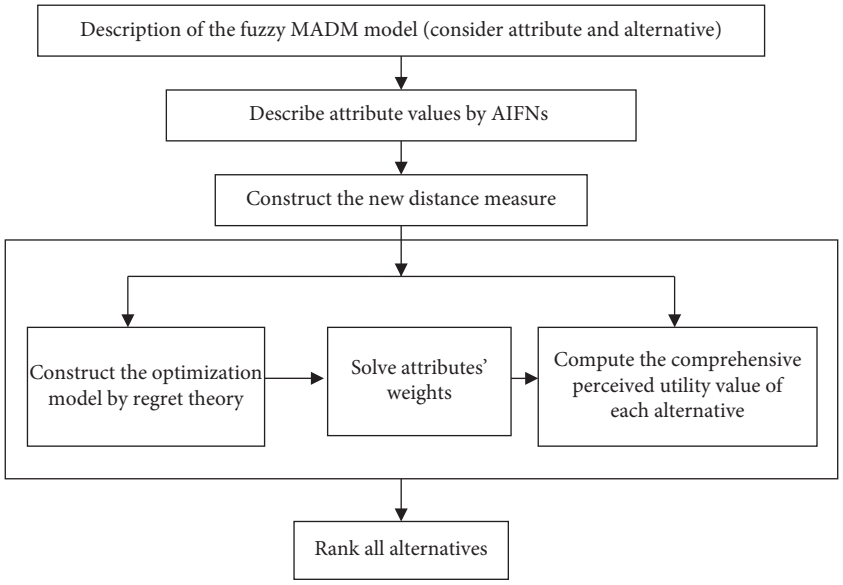


Figure 2: The calculation process of the intuitionistic fuzzy MADM model.

The negative ideal point is defined as $x^- = (x_1^-, x_2^-, \ldots, x_n^-)$, where $x_j^- = <1, 0>$, $j = 1, 2, \ldots, n$

Step 3: calculate the distances $d(x_{ij}, x_j^+)$ and $d(x_{ij}, x_j^-)$, where $d(A, B)$ represents the intuitionistic fuzzy Canberra distance between $A = <\mu_A, \nu_A>$ and $B = <\mu_B, \nu_B>$. Then,

$$d\left(x_{ij}, x_j^+\right) = \frac{1-\mu_{ij}}{3+\mu_{ij}} + \frac{\nu_{ij}}{2+\nu_{ij}} + \frac{1-\mu_{ij}-\nu_{ij}}{3-\mu_{ij}-\nu_{ij}},$$
$$d\left(x_{ij}, x_j^-\right) = \frac{\mu_{ij}}{2+\mu_{ij}} + \frac{1-\nu_{ij}}{3+\nu_{ij}} + \frac{1-\mu_{ij}-\nu_{ij}}{3-\mu_{ij}-\nu_{ij}}. \tag{19}$$

Step 4: according to equation (14), calculate the perceived utility function value F_{ij} of attribute value x_{ij} corresponding to each alternative.

Step 5: establish optimization model (17), and calculate the optimal weight vector $\mathbf{w}^*$ with the help of MATLAB software.

Step 6: the optimal weight obtained from Step 5 is substituted into equation (18), and the comprehensive perceived utility value of each alternative is obtained. The merits and demerits of the scheme are determined according to the comprehensive perceived utility value H_i. The higher the value of H_i, the better the corresponding alternative x_i.

5. Numerical Example

The effectiveness and practicability of this method are illustrated by an example of assembly parts' supplier selection in Xu [39]. With the economic globalization and the continuous expansion of enterprise scale, the problem of supplier selection has become an important management decision-making problem that all large enterprises need to seriously consider. Let a manufacturing company prepare to find the best supplier in the world for purchasing the most critical parts in the assembly process.

After the primary selection, there are five alternative suppliers x_i $(i = 1, 2, 3, 4, 5)$. Now, the company will evaluate the suppliers according to the following five evaluation indicators (attributes): product price (o_1), product quality (o_2), service performance (o_3), supplier's situation (o_4), and risk factors (o_5). After experts' discussion, the evaluation values of each attribute of the candidate suppliers are finally obtained. Suppose the manufacturing company invites N experts to make the judgment. They are expected to answer "Yes" or "No" or "I do not know" to the question whether alternative x_i satisfies attribute o_j. Let $Q_Y(i, j)$ and $Q_N(i, j)$ denote the sum of "Yes" and "No," respectively. Then, the degrees to which alternative A_i satisfies and does not satisfy attribute o_j can be calculated as

$$\mu_{ij} = \frac{Q_Y(i, j)}{N},$$
$$\nu_{ij} = \frac{Q_N(i, j)}{N}. \tag{20}$$

Then, the evaluation values are expressed by AIFNs, as shown in Table 1.

It is assumed that the attribute weight information is partially known, and the attribute weight satisfies

$$H = \left\{ w = (w_1, w_2, \ldots, w_5)^T \,\middle|\, w_1 \le 0.3, 0.1 \le w_2 \le 0.2, 0.2 \le w_3 \le 0.5, 0.1 \le w_4 \le 0.3 w_3 - w_1 \le 0.1, w_4 \ge w_1, w_5 \le 0.4, w_3 - w_2 \ge w_5 - w_4 \right\}. \tag{22}$$

Table 1: Attribute evaluation value of each alternative supplier o_1.

Suppliers	Evaluation attribute				
	o_1	o_2	o_3	o_4	o_5
x_1	<0.449, 0.370>	<0.565, 0.162>	<0.705, 0.232>	<0.730, 0.170>	<0.646, 0.354>
x_2	<0.719, 0.188>	<0.630, 0.232>	<0.448, 0.378>	<0.557, 0.160>	<0.597, 0.192>
x_3	<0.546, 0.192>	<0.727, 0.182>	<0.641, 0.322>	<0.399, 0.200>	<0.658, 0.192>
x_4	<0.520, 0.337>	<0.630, 0.100>	<0.539, 0.271>	<0.679, 0.188>	<0.708, 0.198>
x_5	<0.727, 0.128>	<0.520, 0.299>	<0.619, 0.318>	<0.618, 0.229>	<0.609, 0.120>

Next, we use the proposed decision-making method to sort the five suppliers and choose the best desirable supplier.

These suppliers are sorted according to $H_i\,(i = 1, 2, 3, 4, 5)$ from large to small. The result is $x_4 \succ x_5 \succ x_1 \succ x_3 \succ x_2$, and supplier x_4 is the best choice.

Step 1: calculate the score $S_{ij} = \mu_{ij} - \nu_{ij}$ of attribute $x_{ij} = <\mu_{ij}, \nu_{ij}>$. Then, we get the score matrix $S = (S_{ij})_{5\times5}$, and the calculation result is shown in Table 2.

Step 2: determine the positive and negative ideal point.

The positive ideal point is defined as

$$x^+ = \left(x_1^+, x_2^+, \ldots, x_5^+\right), \tag{23}$$

where $x_j^+ = <1, 0>,\ j = 1, 2, \ldots, 5$.

The negative ideal point is defined as

$$x^- = \left(x_1^-, x_2^-, \ldots, x_5^-\right), \tag{24}$$

where $x_j^- = <0, 1>,\ j = 1, 2, \ldots, 5$.

Step 3: calculate the Canberra distances $d\left(x_{ij}, x_j^+\right)$ and $d\left(x_{ij}, x_j^-\right)$, and the results are shown in Tables 3 and 4.

Step 4: the perceived utility function F_{ij} of attribute x_{ij} corresponding to each alternative is calculated. In this paper, $\alpha = 0.88$ and $\delta = 0.3$ are used for calculation, and the calculation results are shown in Table 5.

Step 5: according to equation (17), the following linear programming model is established:

$$\begin{aligned} \max V &= 1.8739w_1 + 2.2942w_2 + 1.5653w_3 \\ &\quad + 2.2187w_4 + 2.3540w_5, \\ \text{s.t.} &\begin{cases} w_1 \le 0.3, \\ 0.1 \le w_2 \le 0.2, \\ 0.2 \le w_3 \le 0.5, \\ 0.1 \le w_4 \le 0.3, \\ w_3 - w_1 \le 0, \\ w_4 \ge w_1, \\ w_5 \le 0.4, \\ w_3 - w_2 \ge w_5 - w_4, \\ w_1 + w_2 + w_3 + w_4 + w_5 = 1, \\ w_1, w_2, \ldots, w_5 \ge 0. \end{cases} \end{aligned} \tag{25}$$

Table 2: Score matrix of attribute values of each alternative under each attribute.

Suppliers	Evaluation attribute				
	o_1	o_2	o_3	o_4	o_5
x_1	0.0790	0.4030	0.4730	0.5600	0.2920
x_2	0.5310	0.3980	0.0700	0.3970	0.4050
x_3	0.3540	0.5450	0.3190	0.1990	0.4660
x_4	0.1830	0.5300	0.2680	0.4910	0.5100
x_5	0.5990	0.2210	0.3010	0.3900	0.4890

Table 3: Distance set between attribute values of each scheme and corresponding values of PIS.

	x_1^+	x_2^+	x_3^+	x_4^+	x_5^+
x_1	0.3989	0.3171	0.2141	0.1983	0.2475
x_2	0.2059	0.2704	0.3991	0.3226	0.2951
x_3	0.3314	0.2002	0.2554	0.4347	0.2509
x_4	0.3473	0.2685	0.3364	0.2355	0.2137
x_5	0.2010	0.3494	0.2730	0.2786	0.2843

Table 4: Distance set between attribute values of each scheme and corresponding values of NIS.

	x_1^+	x_2^+	x_3^+	x_4^+	x_5^+
x_1	0.4533	0.6054	0.5288	0.5768	0.4367
x_2	0.5636	0.5417	0.4472	0.6076	0.5784
x_3	0.5834	0.5672	0.4650	0.5833	0.5705
x_4	0.4718	0.6488	0.5219	0.5705	0.5571
x_5	0.6130	0.5018	0.4724	0.5458	0.6348

Table 5: Perceived utility function values of attribute values corresponding to each alternative.

Suppliers	Evaluation attribute				
	o_1	o_2	o_3	o_4	o_5
x_1	0.0739	0.3923	0.5591	0.6703	0.2657
x_2	0.6186	0.4826	0.0060	0.3975	0.4545
x_3	0.3341	0.6254	0.4041	0.1804	0.5028
x_4	0.1167	0.5838	0.3110	0.6002	0.5441
x_5	0.7307	0.2101	0.2850	0.3702	0.5868

Solving model (25), we get the optimal attribute weight vector $\mathbf{w}^* = (0.1, 0.1, 0.2, 0.25, 0.35)^T$.

Step 6: calculate the comprehensive perceived utility value of each alternative, and get

$$\begin{aligned} H_1 &= 0.4190, \\ H_2 &= 0.3698, \\ H_3 &= 0.3979, \\ H_4 &= 0.4728, \\ H_5 &= 0.4490. \end{aligned} \tag{26}$$

According to the result of Xu [39], the ranking order is $x_5 \succ x_4 \succ x_1 \succ x_2 \succ x_3$, which is different from the result of this paper. This is because the optimization model established by the score function in Xu [39] does not consider the influence of the degree of hesitation on the ranking of IFSs.

Most of the existing intuitionistic fuzzy MADM methods are based on the expected utility theory assuming that decision makers are completely rational. However, in the actual MADM process, decision makers often have subjective risk preferences, such as psychological and behavioral factors for alternatives. So, it is important to consider the risk attitude of decision makers in the decision process. In this article, a new regret theory-based decision-making method is proposed for the MADM problem in which attribute values are expressed by AIFNs.

The advantage and limitation of the proposed decision-making method can be summarized as follows:

(1) The comprehensive perceived utility value constructed in this paper not only considers the score function but also considers the decision maker's regret gratification value; therefore, it is more in line with the objective reality

(2) How to determine the most suitable values of parameters in the regret theory is its limitation

The novelty of this paper is the proposition of a new decision-making method based on the regret theory, which can better reflect the psychological and behavioral factors of the decision maker than many existing decision-making methods. This paper also develops a new weighting method based on the intuitionistic fuzzy Canberra distance.

6. Conclusion

For the MADM problem in which attribute values are expressed by AIFNs, this paper develops a new decision-making method based on the regret theory combined with an extension of the Canberra distance measure. The main contributions of this article are as follows:

(1) Regret theory can describe humans' psychological behavior under uncertain conditions more truthfully, and it can explain the phenomena that expected utility theory cannot. Our decision-making method considers the psychological factors of decision makers based on the regret theory, which can be more in line with the reality.

(2) This article first constructs a new distance of IFSs based on the traditional Canberra distance. Then, a new weighting method is put forward by establishing an optimization model, which is a model of the maximum optimal comprehensive perceived utility value under given weighting information. The new method enriches and develops the weight attribute determination method.

(3) A numerical example of supplier selection is utilized to show the effectiveness and feasibility of the proposed method. The proposed MADM method based on the regret theory has advantages of simple calculation process and easy software implementation.

In future, we will apply the proposed MADM method to solve other decision-making problems, such as the risk evaluation, system optimization, and material selection. Furthermore, we will develop the new intuitionistic fuzzy distance for the application of clustering analysis and image processing.

Data Availability

The data supporting this numerical example are from [39] which has been cited.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Acknowledgments

This research was supported by the National Natural Science Foundation of China (no.71661012). This paper was selected as an excellent paper from the 6th International Conference on Fuzzy Systems and Data Mining (FSDM2020) which is to be held online during November 13–16, 2020.

References

[1] X. Xie, D. Yue, H. Zhang et al., "Fault estimation observer design for discrete-time takagi–sugeno fuzzy systems based on homogenous polynomially parameter-dependent Lyapunov functions," *IEEE Transactions on Cybernetics*, vol. 47, no. 9, pp. 2504–2513, 2017.

[2] J. Qin, W. Fu, H. Gao et al., "Distributed k-means algorithm and fuzzy c-means algorithm for sensor networks based on multiagent consensus theory," *IEEE Transactions on Cybernetics*, vol. 47, no. 3, pp. 772–783, 2017.

[3] Y. Z. Han and Y. Deng, "An enhanced fuzzy evidential DEMATEL method with its application to identify critical success factors," *Soft Computing*, vol. 22, pp. 5073–5090, 2018.

[4] G. C. Mahata, "A production-inventory model with imperfect production process and partial backlogging under learning considerations in fuzzy random environments," *Journal of Intelligent Manufacturing*, vol. 28, no. 4, pp. 883–897, 2017.

[5] K. Mittal, P. C. Tewari, and D. Khanduja, "Productivity improvement under manufacturing environment using shainin system and fuzzy analytical hierarchy process: a case study," *The International Journal of Advanced Manufacturing Technology*, vol. 92, pp. 407–421, 2017.

[6] A. Si, S. Das, and S. Kar, "An approach to rank picture fuzzy numbers for decision making problems," *Decision Making: Applications in Management and Engineering*, vol. 2, no. 2, pp. 54–64, 2019.

[7] M. Stankovic, P. Gladovic, and V. Popovic, "Determining the importance of the criteria of traffic accessibility using fuzzy AHP and rough AHP method," *Decision Making: Applications in Management and Engineering*, vol. 2, no. 1, pp. 86–104, 2019.

[8] K. T. Atanassov, "Intuitionistic fuzzy sets," *Fuzzy Sets and Systems*, vol. 20, no. 1, pp. 87–96, 1986.

[9] W. Jiang, B. Y. Wei, X. Liu et al., "Intuitionistic fuzzy power aggregation operator based on entropy and its application in decision making," *International Journal of Intelligent Systems*, vol. 33, no. 1, pp. 49–67, 2018.

[10] I. Montes, N. R. Pal, and S. Montes, "Entropy measures for atanassov intuitionistic fuzzy sets based on divergence," *Soft Computing*, vol. 22, pp. 5051–5071, 2018.

[11] R. Joshi and S. Kumar, "A novel fuzzy decision-making method using entropy weights-based correlation coefficients under intuitionistic fuzzy environment," *International Journal of Fuzzy Systems*, vol. 21, no. 1, pp. 232–242, 2019.

[12] P. Rani, D. Jain, and D. S. Hooda, "Extension of intuitionistic fuzzy TODIM technique for multi-criteria decision making method based on shapley weighted divergence measure," *Granular Computing*, vol. 4, pp. 407–420, 2019.

[13] M. X. Luo and R. R. Zhao, "A distance measure between intuitionistic fuzzy sets and its application in medical diagnosis," *Artificial Intelligence in Medicine*, vol. 89, pp. 34–39, 2018.

[14] N. X. Thao, M. Ali, and F. Smarandache, "An intuitionistic fuzzy clustering algorithm based on a new correlation coefficient with application in medical diagnosis," *Journal of Intelligent & Fuzzy Systems*, vol. 36, no. 1, pp. 189–198, 2019.

[15] F. Zhao, H. Liu, J. Fan et al., "Intuitionistic fuzzy set approach to multi-objective evolutionary clustering with multiple spatial information for image segmentation," *Neurocomputing*, vol. 312, pp. 296–309, 2018.

[16] D. Kumar, R. K. Agrawal, and H. Verma, "Kernel intuitionistic fuzzy entropy clustering for MRI image segmentation," *Soft Computing, Soft Computing*, vol. 24, pp. 4003–4026, 2020.

[17] X. Luo, W. M. Li, and W. Zhao, "Intuitive distance for intuitionistic fuzzy sets with applications in pattern recognition," *Applied Intelligence*, vol. 48, pp. 2792–2808, 2018.

[18] H. Verma, A. Gupta, and D. Kumar, "A modified intuitionistic fuzzy c-means algorithm incorporating hesitation degree," *Pattern Recognition Letters*, vol. 122, pp. 45–52, 2019.

[19] M. Kumar, "Applying weakest t-norm based approximate intuitionistic fuzzy arithmetic operations on different types of intuitionistic fuzzy numbers to evaluate reliability of PCBA fault," *Applied Soft Computing*, vol. 23, pp. 387–406, 2014.

[20] S. Kabir, T. K. Geok, M. Kumar et al., "A method for temporal fault tree analysis using intuitionistic fuzzy set and expert elicitation," *IEEE Access*, vol. 8, pp. 980–996, 2020.

[21] H. K. Sharma, K. Kumari, and S. Kar, "A rough set approach for forecasting models," *Decision Making: Applications in Management and Engineering*, vol. 3, no. 1, pp. 1–21, 2020.

[22] D. E. Bell, "Regret in decision making under uncertainty," *Operations Research*, vol. 30, no. 5, pp. 961–981, 1982.

[23] G. Loomes and R. Sugden, "Regret theory: an alternative theory of rational choice under uncertainty," *The Economic Journal*, vol. 92, no. 368, pp. 805–824, 1982.

[24] X. D. Peng and J. G. Dai, "Approaches to pythagorean fuzzy stochastic multi-criteria decision making based on prospect theory and regret theory with new distance measure and score function," *International Journal of Intelligent Systems*, vol. 32, no. 11, pp. 1187–1214, 2017.

[25] Y. Yang and J. Q. . Wang, "SMAA-based model for decision aiding using regret theory in discrete Z-number context," *Applied Soft Computing*, vol. 65, pp. 590–602, 2018.

[26] M. M. Xia, "A hesitant fuzzy linguistic multi-criteria decision-making approach based on regret theory," *International Journal of Fuzzy Systems*, vol. 20, pp. 2135–2143, 2018.

[27] J. Somasundaram and E. Diecidue, "Regret theory and risk attitudes," *Journal of Risk and Uncertainty*, vol. 55, pp. 147–175, 2017.

[28] C. G. Chorus, "Regret theory based route choices and traffic equilibria," *Transportmetrica*, vol. 8, no. 4, pp. 291–305, 2010.

[29] G. Qu, T. Li, W. Qu et al., "Algorithms for regret theory and group satisfaction degree under interval-valued dual hesitant fuzzy sets in stochastic multiple attribute decision making method," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 3, pp. 3639–3653, 2019.

[30] X. Zhang, Z. P. Fan, and F. D. Chen, "Risky multiple attribute decision making with regret aversion," *Journal of Systems & Management*, vol. 23, no. 1, pp. 111–117, 2014.

[31] H. D. Wang, X. H. Pan, J. Yan et al., "A projection-based regret theory method for multi-attribute decision making under interval type-2 fuzzy sets environment," *Information Sciences*, vol. 512, pp. 108–122, 2020.

[32] M. Kumar, "Intuitionistic fuzzy measures of correlation coefficient of intuitionistic fuzzy numbers under weakest triangular norm," *International Journal of Fuzzy System Applications*, vol. 8, no. 1, pp. 48–64, 2019.

[33] P. Grzegorzewski, "Distances and orderings in a family of intuitionistic fuzzy numbers," in *EUSFLAT Conference*, IEEE, Zittau, Germany, September 2003.

[34] M. Kumar, "Evaluation of the intuitionistic fuzzy importance of attributes based on the correlation coefficient under weakest triangular norm and application to the hotel services," *Journal of Intelligent and Fuzzy Systems*, vol. 36, no. 4, pp. 3211–3223, 2019.

[35] V. Perlibakas, "Distance measures for PCA-based face recognition," *Pattern Recognition Letters*, vol. 25, no. 6, pp. 711–724, 2004.

[36] C. Ananth, M. Karthikeyan, N. Mohananthini, and G. Yamuna, "Adaptive and robust multiple image watermarking using canberra distance and dual tree complex wavelet transform," *Journal of Computational and Theoretical Nanoscience*, vol. 16, pp. 1234–1240, 2019.

[37] S. M. Emran and N. Ye, "Robustness of chi-square and canberra distance metrics for computer intrusion detection," *Quality and Reliability Engineering Internatioal*, vol. 18, no. 1, pp. 19–28, 2002.

[38] Y. M. Yang, R. Jia, H. Xun et al., "Determining the number of instars in simulium quinquestriatum (diptera: simuliidae) using k-means clustering via the canberra distance," *Journal of Medical Entomology*, vol. 55, no. 4, pp. 808–816, 2018.

[39] Z. S. Xu, "Multi-person multi-attribute decision making models under intuitionistic fuzzy environment," *Fuzzy Optimization and Decision Making*, vol. 6, no. 3, pp. 221–236, 2007.

Hindawi
Discrete Dynamics in Nature and Society
Volume 2020, Article ID 3864578, 14 pages
https://doi.org/10.1155/2020/3864578

Research Article

A New Theoretical Interpretation of Measurement Error and Its Uncertainty

Huisheng Shi,[1] Xiaoming Ye,[2] Cheng Xing,[2] and Shijun Ding[2]

[1]*AVIC Changcheng Institute of Metrology & Measurement, Beijing 100095, China*
[2]*School of Geodesy and Geomatics, Wuhan University, Wuhan 430079, China*

Correspondence should be addressed to Xiaoming Ye; xmye@sgg.whu.edu.cn

Received 21 July 2020; Revised 21 August 2020; Accepted 7 September 2020; Published 17 September 2020

Academic Editor: Stefania Tomasiello

The traditional measurement theory interprets the variance as the dispersion of a measured value, which is actually contrary to a general mathematical concept that the variance of a constant is 0. This paper will fully demonstrate that the variance in measurement theory is actually the evaluation of probability interval of an error instead of the dispersion of a measured value, point out the key point of mistake in the traditional interpretation, and fully interpret a series of changes in conceptual logic and processing method brought about by this new concept.

1. Introduction

Human scientific research begins with the measurement of various physical quantities, and measurement is the basis of modern scientific research. Especially for the research of artificial intelligence, it is an essential process to obtain natural information and evaluate its authenticity, which itself is the research area of measurement. Therefore, a rigorous measurement theory should be an important part of the modern scientific theory system. This paper points out the main problems existing in the traditional measurement theory and gives the correct interpretation method for the measurement theory.

In traditional measurement theory, the measured value (or observed value) is considered as a random variable, and variance is interpreted as the dispersion of the measured value, both the precision and uncertainty are defined as the dispersion concept of measured value (or observed value) [1–5], so that people can hardly make clear the conceptual difference between them.

However, in any measurement, both the measured value and every observed value are numerical values. According to probability theory, the variance of a numerical value (constant) is zero. So, how does a numerical value show dispersion? Next, we illustrate the contradictory expression of "variance" in traditional theory.

For example, the measured value of Mount Everest elevation in 2005 is $x = 8844.43\,\mathrm{m}$, and its precision is $\sigma(x) = 0.21\,\mathrm{m}$. But in fact, this mathematical expression gives an equation $\sigma(8844.43\,\mathrm{m}) = 0.21\,\mathrm{m}$, which violates the basic mathematical concept, because the equation $x = 8844.43\,\mathrm{m}$ inevitably leads to the equation $\sigma(x) = \sigma(8844.43)$, and according to the concept of $\sigma^2(C) = 0$ in probability theory, there must be $\sigma(8844.43) = 0$. Although many other measured values $x_1, x_2, \ldots$, can be obtained by repeatedly measuring the height of the Mount Everest, and there can be $x \neq x_1 \neq x_2 \neq \cdots$, the equations $x = 8844.43\,\mathrm{m}$ and $\sigma(x) = \sigma(8844.43)$ still exist. Therefore, the equation $\sigma(x) = \sigma(8844.43) = 0.21$ can never be consistent with mathematical concepts.

It can be seen that in traditional theories, the definition of precision and uncertainty is the dispersion of all possible measured values, but their mathematical expression is the dispersion of a single measured value (a single numerical value). This approach is actually a stealth change of concept and will inevitably lead to a series of conceptual logic troubles.

A question arises. All measured values that diverge from each other can definitely be used to describe a random variable, but a measured value must be a numerical value and belong to a constant. So, how should the measurement theory be interpreted?

So far, in the measurement industry, there is no literature questioning the conceptual category of measured values, such as the recent literature [6].

In references [7–10], the authors proposed some new concepts to reinterpret measurement theory. Ye et al. [7] proposed a new error epistemology that all errors follow a random distribution and cannot be classified as systematic error and random error. Ye et al. [8] pointed out that the standard deviation (variance) is the evaluation value of the probability interval of error, any error has a variance for evaluating its uncertainty, and so on. Ye et al. [9] pointed out that the dispersion and deviation of repeated observations are determined by the changing rules of repeated measurement conditions and it is possible and correct to handle errors according to the function model or the random model. Ye and Ding [10] pointed out that the measured value is a numerical value whose variance is zero and the dispersion of a measured value is an incorrect concept. According to these new concepts, the mathematical expression of the Mount Everest elevation case should be $x = 8844.43$ m and $\sigma(\Delta x) = 0.21$ m, where x represents the measured value and Δx represents its error.

Although these new concepts have been proposed to reinterpret measurement theory, the root of these concepts and the interpretation process have not been fully described mathematically. Therefore, in this paper, the authors follow strict mathematical concept to point out the misunderstanding of the traditional concepts, give a clear interpretation for the origin of these new measurement concepts, and systematically explain a series of changes in theoretical logic and mathematical processing.

2. Constant and Random Variable

In probability theory, a constant is a numerical value, such as 100, 150, $x = 100$, $x = 8844.43$, and so on.

Unlike constants, a random variable is an unknown quantity whose actual value cannot be given. Because the random variable is unknown or uncertain, we can only describe the probability range of its value. In order to study its probability range, it is necessary to study the distribution range of all its possible values (sample space), while all possible values refer to the set of test values of random variables under all permitted possible test conditions (the random test does not have the same conditions). Mathematical expectation and variance are the numerical expressions of probability range of the random variable.

For a random variable L with all possible values $\{L_i\}$, $L \in \{L_i\}$, P_i is the probability that each L_i is L (continuous random variables correspond to the probability density function $P(L)$), and its mathematical expectation is defined as follows:

$$E(L) = \sum_{i=1}^{n} P_i L_i, \quad \text{or } E(L) = \int_{-\infty}^{+\infty} LP(L)\mathrm{d}L, \tag{1}$$

and its variance is defined as the dispersion of its possible values $\{L_i\}$:

$$\sigma^2(L) = E[L - E(L)]^2. \tag{2}$$

This means that the random variable L exists within a probability interval with mathematical expectation $E(L)$ and variance $\sigma^2(L)$ or that mathematical expectation and variance are the evaluation values of its probability interval. Note that describing a random variable requires two parameters: mathematical expectation and variance, both of which are indispensable.

Now, suppose that there is a constant C, and $E(L) = C$ and $\sigma^2(L) = 0$, then

$$E[L - E(L)]^2 = 0. \tag{3}$$

By substituting $E(L) = C$, we obtain $E(L - C)^2 = 0$. Therefore, $L = C$.

That is, when the variance of a random variable L is reduced to zero, it becomes a constant C. In other words, for a constant C, because all its possible values are itself, we obtain

$$E(C) = C, \tag{4}$$

$$\begin{aligned}\sigma^2(C) &= E[C - E(C)]^2 \\ &= 0.\end{aligned} \tag{5}$$

That is to say, a constant is a special random variable, and its mathematical expectation and variance are itself and 0, respectively. Of course, a constant is a known quantity and usually does not need to be expressed in terms of probability.

It can be seen that both constant and random variable have their own mathematical expectation and variance. Therefore, if we can give the variance of a quantity but cannot give its mathematical expectation, there must be a conceptual mistake.

It should be noted that for the random variable $L \in \{L_i\}$, its basic feature is that its value is unknown, but for a sample $L_k \in \{L_i\}$, because it is a numerical value, it is still a constant rather than a random variable, and $E(L_k) = L_k$ and $\sigma^2(L_k) = 0$. Obviously, $E(L) \neq E(L_k)$ and $\sigma^2(L) \neq \sigma^2(L_k)$. That is, constant and random variable are distinguished by whether they have a numerical value, and the sample L_k is a numerical value, which is a constant and cannot be described by the entire set $\{L_i\}$. The conceptual differences between random variable and sample are shown in Table 1.

Example 1. A dice has six faces corresponding to the values 1, 2, 3, 4, 5, and 6, respectively. After the dice is thrown, we do not know its display value, which is a random variable $L \in \{1, 2, 3, 4, 5, 6\}$. Then, according to definitions (1) and (2), $E(L) = 3.5$ and $\sigma^2(L) = 2.92$. Its meaning is that although the display value L is unknown, it exists within a probability interval with 3.5 as the center and 2.92 as the width evaluation. Obviously,

Table 1: The conceptual differences between the random variable and the sample.

Random variable L	Sample L_k
$L \in \{L_i\}$	$L_k \in \{L_i\}$
The value of L is unknown	L_k is a numerical value, which is a constant
It has many possible values	All its possible values are itself
Every sample L_i in $\{L_i\}$ is a possible value of L, that is, $P_i \neq 0\%$	Any other samples L_j in $\{L_i\}$ are definitely not L_k, that is, $P_j = 0\%$ and $P_k = 100\%$
We can only use $\{L_i\}$ to describe its probability range	It is a numerical value, and no probability range needs to be described
$E(L) = \sum P_i L_i$, $\sigma^2(L) = E[L - E(L)]^2$	$E(L_k) = L_k$, $\sigma^2(L_k) = 0$
$E(L) \neq E(L_k)$, $\sigma^2(L) \neq \sigma^2(L_k)$	

$$\left.\begin{aligned} E(1) &= 1 \\ E(2) &= 2 \\ &\vdots \\ E(6) &= 6 \end{aligned}\right\} \neq E(L) = 3.5,$$
$$\left.\begin{aligned} \sigma^2(1) &= 0 \\ \sigma^2(2) &= 0 \\ &\vdots \\ \sigma^2(6) &= 0 \end{aligned}\right\} \neq \sigma^2(L) = 2.92. \tag{6}$$

Example 2. The exam scores of all students in a school are $\{x_i\}$, in which the exam score of a student A is x_0. It is reasonable to express an unknown score x with the mathematical expectation $E(x)$ and variance $\sigma^2(x)$ of $\{x_i\}$, but it is illogical to impose $\sigma^2(x)$ on x_0 because x_0 is a known constant and has no need to be expressed by variance and mathematical expectation. Moreover, $\sigma^2(x)$ is obviously not the dispersion of future exam scores of student A. That is, $E(x) \neq E(x_0)$ and $\sigma^2(x) \neq \sigma^2(x_0)$.

Example 3. Someone's salary is $x_0 = 10000$ RMB, and the salaries of all the employees in his company form a sample sequence $\{x_i\}$. By making the statistics of $\{x_i\}$, we can obtain the mathematical expectation $E(x)$ and variance $\sigma^2(x)$. In exactly the same way, although $x_0 \in \{x_i\}$, we cannot force x_0 to belong to a random variable at all because $x_0 = 10000$ RMB is known.

That is to say, it is reasonable to evaluate the probability of an unknown event with the statistic values of a group of known events, but it is illogical to use it to evaluate the "probability" of a known event.

Now, we suppose there is a random variable L with mathematical expectation $E(L) = C$ and $\Delta C = L - E(L)$, and then, there is

$$\begin{aligned} L &= E(L) + L - E(L) \\ &= C + \Delta C. \end{aligned} \tag{7}$$

For the constant C, there are

$$\begin{aligned} E(C) &= C, \\ \sigma^2(C) &= 0. \end{aligned} \tag{8}$$

For the random variable $\Delta C = L - E(L)$, there are

$$\begin{aligned} E(\Delta C) &= E[L - E(L)] \\ &= E(L) - E(L) \\ &= 0, \end{aligned} \tag{9}$$

$$\begin{aligned} \sigma^2(\Delta C) &= E[\Delta C - E(\Delta C)]^2 \\ &= E\left(\Delta C^2\right) \\ &= E[L - E(L)]^2 \\ &= \sigma^2(L). \end{aligned} \tag{10}$$

This is to say, a random variable L with the mathematical expectation C can be viewed as the superposition of a constant C and a random variable ΔC taking 0 as its mathematical expectation. And it should be noted that the variance $\sigma^2(L)$ or $\sigma^2(\Delta C)$ always has nothing to do with constant C.

3. Origin of Conceptual Troubles in Traditional Theory

Figure 1 shows the schematic diagram of the measurement concept in traditional measurement theory. Because people notice that a measured value is in a state of random change in repeated measurement, the measured value and the random error are considered as random variables, and the variance is the dispersion of the measured value or random error. Besides, the systematic error and the true value are constant in repeated measurement, so the systematic error and the true value are considered as constants which have no variances (or the variance is zero). In this way, according to formula (2), $\sigma^2(x) = E[x - E(x)]^2$. Therefore, traditional textbooks [11–13] usually use the form of $\sigma^2(x)$ or σ_x^2 to express the variance. However, these are obviously inconsistent with the meanings of random variables and constants described in Section 2.

Besides, in actual measurement, we always have to give a numerical value x_0 as the final measured value. Therefore, the actual schematic diagram is shown in Figure 2. According to the concepts in Section 2, although the measured value x_0 is a sample within a random distribution, because the measured value x_0 is a numerical value and $\sigma^2(x_0) = 0$ and $E(x_0) = x_0$, it is illogical to replace $\sigma^2(x_0)$ with $\sigma^2(x)$.

In addition, simply replacing $\sigma^2(x_0)$ with $\sigma^2(x)$ cannot express a complete mathematical meaning because the

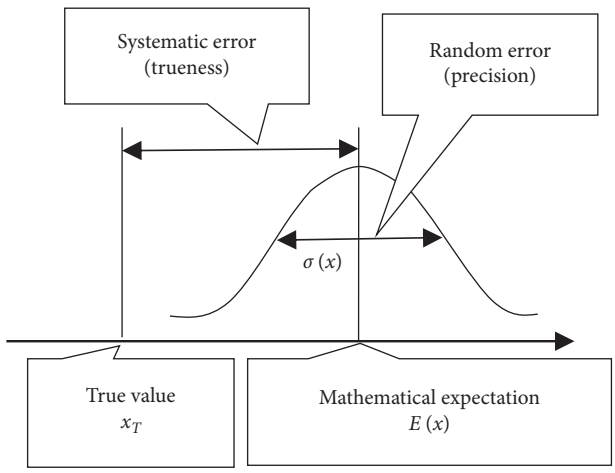


Figure 1: Schematic diagram in traditional measurement theory.

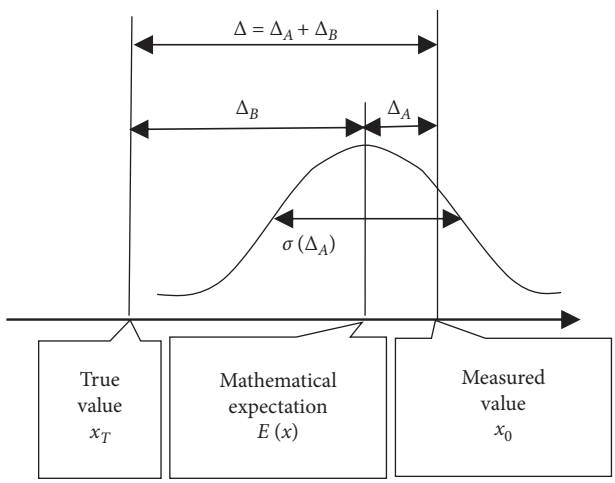


Figure 2: Schematic diagram in the new concept theory.

traditional theory cannot submit the mathematical expectation $E(x)$.

On the contrary, the systematic error and the true value are unknown and are regarded as constants by traditional theory. However, according to formula (4), the mathematical expectation of a constant is itself, so it is impossible to give the numerical values of mathematical expectations of the systematic error and true value. Therefore, this so-called constant is obviously not the same concept as the constant in the probability theory, and it is also a conceptual trouble in the traditional measurement theory.

In short, in the traditional theory, except for the conceptual trouble of violating the concept that the variance of a constant is zero, the conceptual trouble of missing mathematical expectations is shown in Table 2.

Because random errors have variance but systematic errors cannot be quantitatively evaluated, traditional theories believe that trueness and accuracy are both qualitative concepts, and the relationship between the uncertainty concept and them is of course very difficult to explain.

4. Probability Expression of Basic Measurement Concepts

It can be seen that because of the wrong understanding of the concept of the random variable, the conceptual logic of traditional theory actually has systematic troubles. Therefore, we need to reorganize the basic measurement conceptual logic according to the concepts in Section 2.

4.1. Measured Value. In Figure 2, the measured value x_0 is an observed value within all possible observed values $\{x_i\}$ and is a numerical value. According to formulas (4) and (5), there are

$$\begin{aligned} E(x_0) &= x_0, \\ \sigma^2(x_0) &= 0. \end{aligned} \tag{11}$$

Please note that any quantity with a numerical value, including the error sample, the measured value of error, the detected value of instrument error, the value of mathematical expectation, and the value of variance, is a constant.

4.2. Error. As shown in Figure 2, the true value of the measurand is x_T, and the error of the final measured value x_0, which is an unknown deviation $\Delta = x_0 - x_T$, can be divided into $\Delta_A = x_0 - E(x)$ and $\Delta_B = E(x) - x_T$.

First of all, the error Δ_A is a random variable, and $\Delta_A \in \{x_i - E(x)\}$. Moreover, the sample space of error $x - E(x)$ is also $\{x_i - E(x)\}$, so error $x - E(x)$ can be used to represent Δ_A, that is, $\Delta_A = x - E(x)$. According to the formulas (1) and (2), there are

$$\begin{aligned} E(\Delta_A) &= E[x - E(x)] \\ &= 0, \end{aligned} \tag{12}$$

$$\begin{aligned} \sigma^2(\Delta_A) &= E\{x - E(x) - E[x - E(x)]\}^2 \\ &= E[x - E(x)]^2 \\ &= E\left(\Delta_A^2\right). \end{aligned} \tag{13}$$

That is, although deviation Δ_A is unknown, it exists within a probability interval with 0 as center and $\sigma^2(\Delta_A)$ as width evaluation. In other words, $\sigma^2(\Delta_A)$ is the evaluation of probability interval of deviation Δ_A, which expresses the degree that a surveyor cannot determine the value of deviation Δ_A.

Taking the normal distribution as an example, the variance $\sigma^2(\Delta_A)$ expresses that the deviation Δ_A is within the interval of $[-\sigma(\Delta_A) + \sigma(\Delta_A)]$ under the confidence probability of 68%. Variance is actually a concept of error range with probability meaning and expresses an error's possible deviation degree.

In formula (13), the single deviation Δ_A is a member within all its possible values, and the dispersion interval of all its possible values is the probability interval of this deviation Δ_A. An unknown deviation follows a random distribution, which means that all possible values of the deviation follow a random distribution.

Table 2: The conceptual trouble of missing mathematical expectations.

	Measured value x	Random error $x - E(x)$	Systematic error $E(x) - x_T$	True value x_T
Mathematical expectation	Absent	0	Absent	Absent
Variance	$\sigma^2(x)$	$\sigma^2(x)$	0	0

This principle obviously can be extended to the error Δ_B. In fact, when we trace back to the upstream measurement of forming error Δ_B, we will find that the formation principle of error Δ_B is similar to that of current error Δ_A and that the error Δ_B is also a member within all its possible values. Therefore, there is also variance $\sigma^2(\Delta_B)$ to evaluate the probability interval of error Δ_B, and also $E(\Delta_B) = 0$.

For example, the multiplicative constant error R of a geodimeter [14, 15] comes from the frequency error of the quartz crystal and is always viewed as a systematic error without variance by traditional measurement theory. However, it is the output error in the field of instrument manufacturing, and the submission process of its variance will be demonstrated in a case in Section 6.

Obviously, according to the principle of formulas (7)–(10), if the mathematical expectation of an error is C rather than 0, then C must be corrected to the final measured value, and the mathematical expectation of the remaining unknown error is still 0. That is, for any unknown error Δx, there is always

$$E(\Delta x) = 0. \tag{14}$$

Thus, the error's variance is expressed as follows:

$$\begin{aligned} \sigma^2(\Delta x) &= E[x - E(\Delta x)]^2 \\ &= E\left(\Delta x^2\right). \end{aligned} \tag{15}$$

Δx in formulas (14) and (15) can express not only the deviation Δ_A between the measured value and the mathematical expectation, but also the deviation Δ_B between the mathematical expectation and the true value. It can even express the deviation $\Delta = \Delta_A + \Delta_B$ between the measured value and the true value.

In this way, according to formulas (1) and (2), there are

$$\begin{aligned} \Delta &= \Delta_A + \Delta_B, \\ E(\Delta) &= E(\Delta_A) + E(\Delta_B) \\ &= 0, \\ \sigma^2(\Delta) &= E[\Delta - E(\Delta)]^2 \\ &= E\left(\Delta^2\right), \\ &= E(\Delta_A + \Delta_B)^2, \\ &= \sigma^2(\Delta_A) + \sigma^2(\Delta_B). \end{aligned} \tag{16}$$

Because the final measured value is unique and constant, both Δ_A and Δ_B are unknown deviations. In addition, both of them have their own variance; hence, it is incorrect that the traditional measurement theory considers Δ_A as the random error and considers Δ_B as the systematic error. Moreover, the corresponding concepts of precision and trueness are also incorrect.

It should be emphasized that formula (14) means that the mean value of all possible values of an unknown error is 0, which expresses the probabilities that an unknown error takes a positive value and a negative value are equal in our subjective cognition. From a statistical perspective, all possible values of an error refer to the set of all error values under all possible measurement conditions permitted by measurement specification, so the traditional concept of "repeated measurement under the same conditions" must be abandoned [8, 9]; otherwise, a unique error value obtained under a particular condition is only one sample within all possible values and does not represent all possible values, which is very easy to cause the illusion of $E(\Delta x) \neq 0$.

In short, being different from the measured value, the error is unknown; because the error is unknown, we can only study its probability range; because of studying its probability range, we must study all possible values of error; because of studying all possible values of error, error samples must come from all the possible measurement conditions permitted by measurement specification, and the traditional concept of "repeated measurement under the same measurement conditions" must be abandoned.

4.3. Variance of Regular Error. Any error has its variance, including the regular error, because the regular error also has all its possible values.

For example, the periodic error of a phase photoelectric distance meter [14, 15] conforms to the sine regularity, and its function model is $\delta = A\sin((D/\lambda) \times 2\pi + \varphi)$. However, when we only observe the density distribution of all its possible values, this cyclic error's probability density function $f(\delta)$ can be derived as

$$f(\delta) = \begin{cases} \dfrac{1}{\pi\sqrt{A^2 - \delta^2}}, & |\delta| \le A, \\ 0, & |\delta| > A. \end{cases} \tag{17}$$

As shown in Figure 3, furthermore, its variance can be derived as $\sigma^2(\delta) = A^2/2$, and its mathematical expectation can be derived as $E(\delta) = 0$.

Another example, the rounding error δ is a sawtooth cycle regularity function of a true value w. However, when we only observe the density distribution of all possible values, the error also follows a random distribution, as shown in Figure 4, and its probability density function is

$$f(\delta) = \begin{cases} \dfrac{1}{2a}, & |\delta| \le a, \\ 0, & |\delta| > a. \end{cases} \tag{18}$$

Its variance can be derived as $\sigma^2(\delta) = a^2/3$, and its mathematical expectation can be derived as $E(\delta) = 0$.

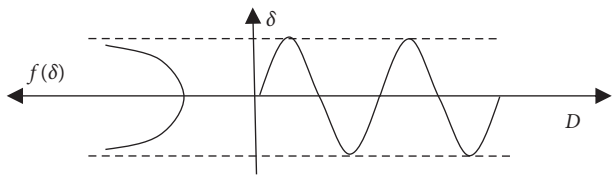


Figure 3: Regularity and randomness of periodic error.

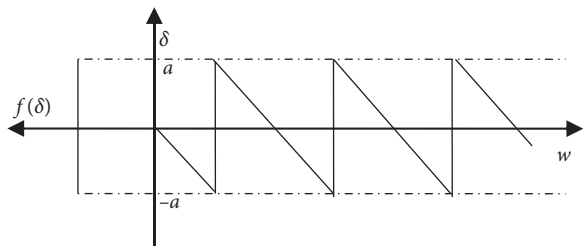


Figure 4: Regularity and randomness of rounding error.

That is to say, the regularity and the randomness are the effect of observing all possible values of error from different perspectives, there is no opposition between them, and there is actually no need to dwell on the error's regularity in the discussion of error evaluation. In other words, when a regular error is unknown, we can still use the mathematical expectation and the variance to describe its probability range. Furthermore, traditional measurement theories use the regularity and the randomness to classify errors into systematic errors and random errors, which is also proved to be incorrect.

In addition, when the measurement conditions vary with the repeated measurement, the corresponding regular errors will cause the dispersion of the repeated observations, which is exactly the same as the dispersion caused by the noise which varies with time conditions [9]. This dispersion is exactly the means by which we obtain its variance. It can be seen that, in order to obtain the variance of an error, we need to collect error samples under all possible measurement conditions.

4.4. Probability Expression of True Value. With previous statements, we already know that for the measured value x_0, there are $E(x_0) = x_0$ and $\sigma^2(x_0) = 0$; for the error Δ, there are $E(\Delta) = 0$ and $\sigma^2(\Delta) = E(\Delta)^2$. Because the error is the difference between the measured value and the true value, that is, $\Delta = x_0 - x_T$, there are

$$\begin{aligned} x_T &= x_0 - \Delta, \\ E(x_T) &= E(x_0 - \Delta) \\ &= E(x_0) - E(\Delta) \\ &= x_0, \\ \sigma^2(x_T) &= E[x_T - E(x_T)]^2 \\ &= E(x_0 - \Delta - x_0)^2 \\ &= \sigma^2(\Delta). \end{aligned} \tag{19}$$

The probability expressions of the true value x_T, the measured value x_0, and the error Δ are summarized in Table 3 [10].

The above is the case where an observed value is used as the final measured value. If the mean value of n observations is taken as the final measured value, it can be inferred that the variance $\sigma^2(\Delta_A)$ will decrease by n times (Section 6.1).

5. Covariance Propagation

It can be seen from Table 3 that after a measured value x_0 is given, only variance $\sigma^2(\Delta)$ needs to be studied.

5.1. Covariance. Considering that two correlated errors have common component, formula (15) can be extended to any two errors. That is,

$$\sigma(\Delta x_i \Delta x_j) = E(\Delta x_i \Delta x_j). \tag{20}$$

Thus, for the error sequence $\Delta\mathbf{X} = (\Delta x_1 \ \ \Delta x_2 \ \cdots \ \Delta x_t)^T$, the definition of variance is

$$\mathbf{D}(\Delta\mathbf{X}) = E(\Delta\mathbf{X})(\Delta\mathbf{X})^T. \tag{21}$$

That is,

$$\begin{aligned} \mathbf{D}(\Delta\mathbf{X}) &= E\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_t \end{pmatrix} (\Delta x_1 \ \ \Delta x_2 \ \cdots \ \Delta x_t) \\ &= \begin{pmatrix} \sigma_{11}^2 & \sigma_{12} & \cdots & \sigma_{1t} \\ \sigma_{21} & \sigma_{22}^2 & \cdots & \sigma_{12} \\ \vdots & \vdots & & \vdots \\ \sigma_{t1} & \sigma_{t2} & \cdots & \sigma_{tt}^2 \end{pmatrix}. \end{aligned} \tag{22}$$

Obviously, the true definition of the variance is formula (21). Formula (15) is only a special case of formula (21) when $t = 1$, and formula (13) is only a special case of formula (15) when interpreting Δx as $x - E(x)$.

So, what is the meaning of covariance?

It is assumed that the errors k, p, and q are uncorrelated with each other, and that their variance are σ_k^2, σ_p^2, and σ_q^2, respectively. Now, there are two errors $\delta = k + p$ and $\varepsilon = k + q$, and they contain a communal error component k. Therefore, we can obtain

$$\begin{aligned} \sigma_\delta^2 &= \sigma_k^2 + \sigma_p^2, \\ \sigma_\varepsilon^2 &= \sigma_k^2 + \sigma_q^2. \end{aligned} \tag{23}$$

According to the definition of covariance,

$$\begin{aligned} \sigma_{\delta\varepsilon} &= E(\delta\varepsilon) \\ &= E[(k+p)(k+q)] \\ &= E(k^2) + E(kp) + E(kq) + E(pq), \end{aligned} \tag{24}$$

Table 3: The probability expression of the true value, measured value, and error.

	Measured value x_0	Error Δ	True value x_T
Mathematical expectation	x_0	0	x_0
Variance	0	$\sigma^2(\Delta)$	$\sigma^2(\Delta)$

with the assumption that the errors k, p, and q are irrelevant from each other, we can obtain $E(kp) = 0$, $E(kq) = 0$, and $E(pq) = 0$, and equation (24) becomes

$$\begin{aligned}\sigma_{\delta\varepsilon} &= E\left(k^2\right)\\ &= \sigma_k^2.\end{aligned} \tag{25}$$

The covariance $\sigma_{\delta\varepsilon}$ is actually the variance of their communal error component k. That is to say, the mathematical meaning of covariance is the probability evaluation of the communal error component contained in two errors. As long as there are communal error components among different errors, there must be a covariance between them. Of course, the symbol and coefficient of communal error component should be considered in the actual measurement.

For example, the two measured value's errors measured by the same instrument are correlated, and the errors of two instruments calibrated by the same benchmark are also correlated.

Moreover, like the above principle, when two errors are associated with the same measurement condition, there is also a covariance between them. For example, both the error of light speed in atmosphere and the thermal expansion error of metal are functions of temperature, and there is a correlation between all possible values of the two errors.

5.2. Law of Covariance Propagation. Because any error has all its possible value and has its variance, the law of covariance propagation is extended to any error. In addition, the law of covariance propagation can only be interpreted as the propagation law of error's probability interval and cannot be interpreted as propagation law of measured value's dispersion.

Here is a measurement equation:

$$\mathbf{Z} = F(\mathbf{X}). \tag{26}$$

We can derive the total differential of equation (26):

$$\Delta\mathbf{Z} = \mathbf{K}\cdot\Delta\mathbf{X}, \tag{27}$$

where

$$\Delta\mathbf{Z} = \begin{pmatrix}\Delta Z_1\\ \Delta Z_2\\ \vdots\\ \Delta Z_t\end{pmatrix},$$

$$\begin{aligned}\mathbf{K} &= \begin{pmatrix}k_{11} & k_{12} & \cdots & k_{1n}\\ k_{21} & k_{22} & \cdots & k_{2n}\\ \vdots & \vdots & \vdots & \vdots\\ k_{t1} & k_{t2} & \cdots & k_{tn}\end{pmatrix},\\ \Delta\mathbf{X} &= \begin{pmatrix}\Delta x_1\\ \Delta x_2\\ \vdots\\ \Delta x_n\end{pmatrix}.\end{aligned} \tag{28}$$

According to formula (21), the covariance matrix of the error sequence $\Delta\mathbf{Z}$ is

$$\begin{aligned}\mathbf{D}(\Delta\mathbf{Z}) &= E(\Delta\mathbf{Z})(\Delta\mathbf{Z})^T\\ &= E(\mathbf{K}\cdot\Delta\mathbf{X})(\mathbf{K}\cdot\Delta\mathbf{X})^T\\ &= \mathbf{K}\cdot\mathbf{D}(\Delta\mathbf{X})\cdot\mathbf{K}^T.\end{aligned} \tag{29}$$

Equation (29) is the law of covariance propagation. The relationship between equations (27) and (29) is as follows:

(1) In the error equation (27), the direct participants of synthesis are the error itself, each error is a deviation, and error synthesis always follows the algebra rule.

(2) In the variance equation (29), the participants of synthesis are all possible values of each error instead of each error itself. It expresses the propagation relation of dispersion of all possible values of errors and also the propagation relation of probability intervals between errors.

6. Statistical Calculation of Variance

Since the number of error samples is always limited in actual measurement, formula (15) can be approximated as

$$\sigma^2(\Delta x) \approx \frac{\sum_{i=1}^{n}(\Delta x_i)^2}{n}. \tag{30}$$

Formula (30) is also the source of least squares principle. That is to say, from the perspective of the new concept, only the concept of error evaluation changes, while the principle of the least square method used to obtain the best measured values does not change.

In actual measurement, in order to achieve the reduction and evaluation of the measurement error, a large number of observations should be carried out. Because errors make a large number of observations contradict from each other, the optimal measured values must be given by adjustment process, and the errors of the measured values should also be evaluated. We only discuss the case of the least square adjustment in this section.

6.1. Direct Measurement for Single Measurand. A measurand is directly measured by n times, and n observations x_i are obtained. In this way, using y_0 to represent the best measured value, the error equations are

$$\left.\begin{aligned} v_1 &= x_1 - y_0 \\ v_2 &= x_2 - y_0 \\ &\vdots \\ v_n &= x_n - y_0 \end{aligned}\right\}. \tag{31}$$

According to the least squares method, the final measured value is

$$y_0 = \sum_{i=1}^{n} \frac{x_i}{n}. \tag{32}$$

The measurement model of this measurement method is $V = X - Y$, and v_i, x_i, and y_0 are the samples of random variables V, X, and Y, respectively.

Taking the total differentiation of equation (32), the error propagation equation is

$$\Delta y = \sum_{i=1}^{n} \frac{\Delta x_i}{n}. \tag{33}$$

Now, we only discuss the variances of the error components $x_i - E(X)$ and $y_0 - E(Y)$ and make $\Delta x_i = x_i - E(X)$ and $\Delta y = y_0 - E(Y)$.

Because $\Delta x_i = x_i - E(X)$ is unknown, is a random variable, and has the same sample space $\{x_i - E(X)\}$ as $X - E(X)$, $\Delta x_i = x_i - E(X)$ can be represented by $\Delta x = X - E(X)$, that is, $\Delta x_i = \Delta x = X - E(X)$. Similarly, $\Delta y = Y - E(Y)$.

By applying the law of covariance propagation to formula (33), there is

$$\sigma^2(\Delta y) = \frac{\sigma^2(\Delta x)}{n}. \tag{34}$$

According to the measurement model $V = X - Y$, there is

$$\begin{aligned} \Delta x &= X - E(X) \\ &= V + Y - E(V + Y) \\ &= V + \Delta y. \end{aligned} \tag{35}$$

Therefore, according to the definition of variance, there is

$$\begin{aligned} \sigma^2(\Delta x) &= E\left[(\Delta x)^2\right] \\ &= E[V + \Delta y]^2 \\ &= E(V)^2 + E(\Delta y)^2 \\ &= E(V)^2 + \sigma^2(\Delta y). \end{aligned} \tag{36}$$

Substituting $E(V)^2 \approx (1/n)\sum_{i=1}^{n} v_i^2$ and formula (34) into equation (36), we obtain

$$\sigma^2(\Delta x) \approx \frac{\sum_{i=1}^{n} v_i^2}{n} + \frac{\sigma^2(\Delta x)}{n}. \tag{37}$$

Therefore,

$$\sigma(\Delta x) \approx \sqrt{\frac{\sum_{i=1}^{n} v_i^2}{n-1}}. \tag{38}$$

6.2. Indirect Measurement for Single Measurand. Different from the direct measurement, each observation x_i in the indirect measurement is the measured data of a_i times of measurand. The error equations of the repeated measurement are

$$\left.\begin{aligned} v_1 &= x_1 - a_1 y_0 \\ v_2 &= x_2 - a_2 y_0 \\ &\vdots \\ v_n &= x_n - a_n y_0 \end{aligned}\right\}. \tag{39}$$

According to the least squares method, the final measured value is

$$y_0 = \frac{\sum_{i=1}^{n} a_i x_i}{\sum_{i=1}^{n} a_i^2}. \tag{40}$$

The measurement model is $V_i = X_i - a_i Y$, and v_i, x_i, and y_0 are the samples of random variables V_i, X_i, and Y, respectively.

Similarly, for the errors $\Delta x = X_i - E(X_i)$ and $\Delta y = Y - E(Y)$, the covariance propagation relationship is

$$\sigma^2(\Delta y) = \frac{\sigma^2(\Delta x)}{\sum_{i=1}^{n} a_i^2}. \tag{41}$$

Similarly, according to the measurement model $V_i = X_i - a_i Y$, there is

$$\begin{aligned} X_i - E(X_i) &= V_i + a_i Y - E(V_i + a_i Y) \\ &= V_i + a_i \Delta y. \end{aligned} \tag{42}$$

According to the definition of variance, there is

$$\begin{aligned} \sigma^2(\Delta x) &= E[\Delta x]^2 \\ &= \lim_{n\to\infty} \frac{1}{n}\left\{[X_1 - E(X_1)]^2 + [X_2 - E(X_2)]^2 + \cdots\right\} \\ &= \lim_{n\to\infty} \frac{1}{n}\left[(V_1 + a_1\Delta y)^2 + (V_2 + a_2\Delta y)^2 + \cdots\right] \\ &\approx \frac{1}{n}\sum_{i=1}^{n} v_i^2 + \frac{\sum_{i=1}^{n} a_i^2}{n}\sigma^2(\Delta y) \\ &= \frac{1}{n}\sum_{i=1}^{n} v_i^2 + \frac{\sigma^2(\Delta x)}{n}. \end{aligned} \tag{43}$$

Therefore,

$$\sigma(\Delta x) \approx \sqrt{\frac{\sum_{i=1}^{n} v_i^2}{n-1}}. \tag{44}$$

6.3. Indirect Measurement for Multiple Measurands. In this measurement mode, there are t different measurands, and each observation value x_i is obtained by measuring the linear superposition value of multiple measurands. The error equations of the repeated measurements are

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} - \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1t} \\ a_{21} & a_{22} & \cdots & a_{2t} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nt} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_t \end{pmatrix}. \quad (45)$$

That is,

$$\mathbf{V} = \mathbf{X} - \mathbf{AY}. \quad (46)$$

According to the principle of the least squares, its measured values are

$$\mathbf{Y} = \left[\mathbf{A}^T\mathbf{A}\right]^{-1}\mathbf{A}^T\mathbf{X}. \quad (47)$$

The measurement model is $V_i = X_i - \left(a_{i1} \;\; a_{i2} \;\; \cdots \;\; a_{\mathrm{it}}\right)\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_t \end{pmatrix}$, and v_i, x_i, and y_j are the samples of random variables V_i, X_i, and Y_j, respectively.

The error propagation equation is

$$\Delta\mathbf{Y} = \left[\mathbf{A}^T\mathbf{A}\right]^{-1}\mathbf{A}^T\Delta\mathbf{X}. \quad (48)$$

Similarly, for the errors $\Delta x = X_i - E(X_i)$ and $\Delta y_j = Y_j - E(Y_j)$, the covariance propagation relationship is

$$\mathbf{D}(\Delta\mathbf{Y}) = \sigma^2(\Delta x)\left[\mathbf{A}^T\mathbf{A}\right]^{-1}. \quad (49)$$

Similarly, according to the measurement model, there is

$$X_i - E(X_i) = V_i + \left(a_{i1} \;\; a_{i2} \;\; \cdots \;\; a_{it}\right)\begin{pmatrix} \Delta y_1 \\ \Delta y_2 \\ \vdots \\ \Delta y_t \end{pmatrix}. \quad (50)$$

According to the definition of variance, there is

$$\begin{aligned} \sigma^2(\Delta x) &= E[\Delta x]^2 \\ &= \lim_{n\longrightarrow\infty}\frac{1}{n}\left\{[X_1 - E(X_1)]^2 + [X_2 - E(X_2)]^2 + \cdots\right\} \\ &= \lim_{n\longrightarrow\infty}\frac{1}{n}\sum_{i=1}^{n}\left[V_i + \left(a_{i1} + a_{i2} + \cdots + a_{\mathrm{it}}\right)\begin{pmatrix} \Delta y_1 \\ \Delta y_2 \\ \vdots \\ \Delta y_t \end{pmatrix}\right]^2 \\ &\approx \frac{1}{n}\sum_{i=1}^{n} v_i^2 + \frac{1}{n}\sum_{i=1}^{n}\left[\left(\%\,\Delta y_1 \;\; \Delta y_2 \;\; \cdots \;\; \Delta y_t\right)\mathbf{A}^T\mathbf{A}\begin{pmatrix} \Delta y_1 \\ \Delta y_2 \\ \vdots \\ \Delta y_t \end{pmatrix}\right]. \end{aligned} \quad (51)$$

Omitting the tedious algebraic calculation process, the final result is

$$\sigma(\Delta x) \approx \sqrt{\frac{\sum_{i=1}^{n} v_i^2}{n-t}}. \quad (52)$$

As you can see, the change in Bessel's formula is that $\sigma(x)$ is written as $\sigma(\Delta x)$, and $\sigma(\Delta x)$ represents the dispersion of all possible values of the deviation $\Delta x_i = x_i - E(X)$. Also, the standard deviation $\sigma(\Delta y)$ or $\sigma(\Delta y_j)$, which is given by formula (34), (41), or (49), is also the evaluation of probability interval of the deviation $\Delta y = y_0 - E(Y)$ or $\Delta y_j = y_j - E(Y_j)$. Obviously, it is incorrect to express Bessel's formula as $\sigma^2(x_k) \approx (1/(n-1))\sum_{j=1}^{n}(x_j - \overline{x})^2$ in the literatures [1, 4] because x_k is a numerical value.

For example, the measured frequency values of a quartz crystal at different temperatures are shown in Table 4, and the nominal value of frequency is $f_0 = 5.000050$ MHz. Now, we need to give a temperature correction model for the frequency error and evaluate the standard deviation of the residual error after correction.

We use the first 4 terms of the Taylor series as the temperature model of the frequency error, that is, $R = a + bT + cT^2 + dT^3$.

In this way, the error equation set is

Table 4: Observed values.

Temperature (°C)	Frequency (MHz)	Error value, $R_i = \Delta f_i / f_0 (1 \times 10^{-6})$
−40°	4.999900	−30
−30°	4.999975	−15
−20°	5.000040	−2
−10°	5.000085	7
0°	5.000115	13
10°	5.000110	12
20°	5.000070	4
30°	5.000035	−3
40°	5.000010	−8
50°	4.999995	−11
60°	4.999995	−11
70°	5.000010	−8
80°	5.000045	−1
90°	5.000125	15
100°	5.000235	37

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix} - \begin{pmatrix} 1 & T_1 & T_1^2 & T_1^3 \\ 1 & T_2 & T_2^2 & T_2^3 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & T_n & T_n^2 & T_n^3 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}. \tag{53}$$

According to the least squares method, there is

$$\begin{pmatrix} n & \sum T_i & \sum T_i^2 & \sum T_i^3 \\ \sum T_i & \sum T_i^2 & \sum T_i^3 & \sum T_i^4 \\ \sum T_i^2 & \sum T_i^3 & \sum T_i^4 & \sum T_i^5 \\ \sum T_i^3 & \sum T_i^4 & \sum T_i^5 & \sum T_i^6 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} \sum R_i \\ \sum R_i T_i \\ \sum R_i T_i^2 \\ \sum R_i T_i^3 \end{pmatrix}. \tag{54}$$

Substituting the values in Table 4 into the above equation, there are

$$\begin{pmatrix} 15 & 450 & 41500 & 292500 \\ 450 & 41500 & 2925000 & 256870000 \\ 41500 & 2925000 & 256870000 & 21952500000 \\ 292500 & 256870000 & 21952500000 & 1983295000000 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix} = \begin{pmatrix} -1 \\ 4610 \\ 304500 \\ 42713000 \end{pmatrix}. \tag{55}$$

Solving the equations, the following equation is obtained:

$$\begin{aligned} a &= 9.983251, \\ b &= -0.013518, \\ c &= -0.018601, \\ d &= 0.000214. \end{aligned} \tag{56}$$

Therefore, the frequency error's function model is fitted as

$$R = 9.983251 - 0.013518T - 0.018601T^2 + 0.000214T^3. \tag{57}$$

Figure 5 is the comparison curve between the model and the actual error. According the formula (52), the standard deviation of residual error is

$$\begin{aligned} \sigma(\Delta R) &= \sqrt{\frac{\sum_{i=1}^{n} v_i^2}{n-4}} \\ &= \pm 2.3 \times 10^{-6}. \end{aligned} \tag{58}$$

Finally, the frequency of quartz crystal is given as follows:

$$f = f_0\left(1 + R \times 10^{-6}\right). \tag{59}$$

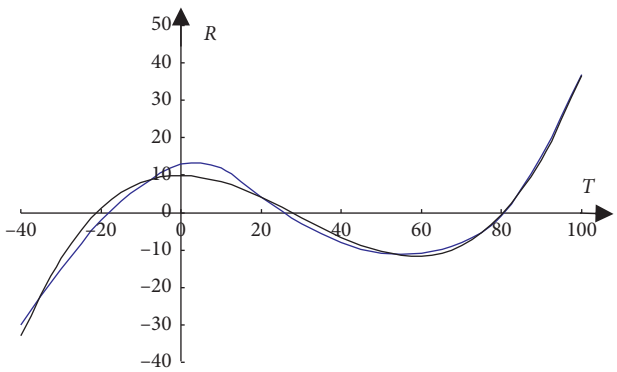


Figure 5: The function model fitting of frequency error.

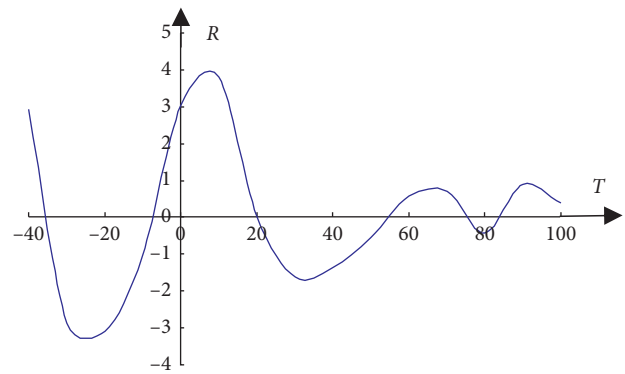


Figure 6: The residual error's curve.

That is, temperature-frequency error can be corrected by the measured value of a temperature sensor, and a more accurate frequency value can be calculated. Residual error (as shown in Figure 6), which is still a regular error, is also processed by statistical rules, and the standard deviation of the residual error is $\pm 2.3 \times 10^{-6}$. This error processing method has been widely used in the manufacture of photoelectric geodimeter [14, 15].

7. Uncertainty

According to Figure 2, the total error of the final measured value is

$$\Delta = \Delta_A + \Delta_B, \tag{60}$$

where Δ_A is the deviation between measured value and expectation and Δ_B is the deviation between expectation and true value.

Because the two errors are usually irrelevant, according to the law of covariance propagation (29), there is

$$\sigma(\Delta) = \sqrt{\sigma^2(\Delta_A) + \sigma^2(\Delta_B)}. \tag{61}$$

This total standard deviation $\sigma(\Delta)$ is the evaluation of probability interval of total error Δ (the dispersion of all possible values of total error Δ). It can be seen that formula (61) is consistent with the traditional uncertainty evaluation (but the expression is changed from $\sigma(x)$ to $\sigma(\Delta)$). Therefore, this total standard deviation $\sigma(\Delta)$ is actually the uncertainty, which expresses the probability range of the total error of final measured value. And the uncertainty concept, which is interpreted as the dispersion of the measured value (constant) in the traditional measurement theory, is also proved to be incorrect.

It can be seen from Table 3 that the uncertainty is also the possible degree that the true value deviates from the measured value. That is, the uncertainty is not only the uncertainty of the error but also the uncertainty of the true value but is not the uncertainty of the measured value. The measured value, which is a certain value, has no uncertainty.

According to the interpretation of the traditional theory, $\sigma(\Delta_A)$ and $\sigma(\Delta_B)$ are referred as the uncertainty of Type A and the uncertainty of Type B, respectively. However, the current $\sigma(\Delta_B)$ is actually $\sigma(\Delta)$ of historical upstream measurement, and the current $\sigma(\Delta)$ can also be used as $\sigma(\Delta_B)$ in future downstream measurement. This kind of interpretation with A/B classification of the uncertainty evaluation is obviously too rigid.

Moreover, currently widely used formula (61) is only applicable to the direct repeated measurement model in Section 6.1 but has no use at all for the indirect repeated measurement in Sections 6.2 and 6.3 because in indirect repeated measurement, there are usually some error sources which not only contribute to dispersion of repeated observations but also contribute to their deviation, and it is difficult to distinguish them with A/B classification method. Therefore, the A/B classification method is not universal in practice.

Formula (61) comes from the covariance propagation law (29). Thus, the basic principle of uncertainty synthesis is covariance propagation law (29), and the uncertainty synthesis does not need to apply the interpretation of A/B classification mechanically. Here is a simple example to illustrate this principle, which is also a comparison with the traditional practice.

For example, four points A, B, C, and D are located on a straight line (Figure 7), and the observation data of distances obtained by geodimeter [14, 15] are shown in Table 5. Please solve the final measured values of each line segment and the uncertainty of each error.

Using y_1, y_2, and y_3 to express the final measured values of AB, BC, and CD, respectively, and using k to express the measured value of the additive constant error of geodimeter, the observation error equation is

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ k \end{pmatrix}. \tag{62}$$

According to the least squares method, there are

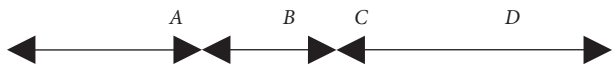


Figure 7: Distance measurement.

Table 5: Observed values.

	Line segment	Observed values
1	AB	$x_1 = 39.8538$ m
2	BC	$x_2 = 159.957$ m
3	CD	$x_3 = 320.0015$ m
4	AC	$x_4 = 199.8117$ m
5	BD	$x_5 = 479.9601$ m
6	AD	$x_6 = 519.8149$ m

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ k \end{pmatrix} = \left[\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}^T\right]^{-1}\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_6 \end{pmatrix}$$
$$= \frac{1}{4}\times\begin{pmatrix} 1 & -2 & -1 & 1 & -1 & 2 \\ -2 & 1 & -2 & 1 & 1 & 1 \\ -1 & -2 & 1 & -1 & 1 & 2 \\ 2 & 2 & 2 & 0 & 0 & -2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_6 \end{pmatrix}. \tag{63}$$

Substituting the numerical values of all observed values into equation (63), we obtain

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ k \end{pmatrix} = \begin{pmatrix} 39.8549 \\ 159.9583 \\ 320.0030 \\ -0.0013 \end{pmatrix}(\text{m}). \tag{64}$$

For the traditional measurement theory, the next step is to substitute (64) into (62), and six residual v_i are obtained. Then $\sigma(x)$ is obtained by using the Bessel formula $\sigma(x) = \sqrt{(\sum_{i=1}^{n} v_i^2)/n-t}$, and $\sigma(y_1)$, $\sigma(y_2)$, $\sigma(y_3)$, and $\sigma(k)$ are obtained by using the covariance propagation law. Finally, $\sigma(y_1)$, $\sigma(y_2)$, $\sigma(y_3)$, and $\sigma(k)$ are called as precision or uncertainty of Type A, but the uncertainty of Type B is almost impossible to discuss.

However, from the perspective of the new conceptual theory, there are three conceptual troubles in the above variance submission process: (1) The degree of the freedom $n-t$ is too small, so it is meaningless to apply the Bessel formula. (2) In Table 5, each observed value x_i is a numerical value, and according to equation (64), each measured value y_j is also a numerical value, so their variances should be 0. (3) The contribution of the covariance between the errors of each observation value x_i has not been taken into account at all (uncertainty synthesis issue).

The following is the variance submission process of the new conceptual theory for this case.

Taking the total differential of equation (63), the error propagation equation is obtained as follows:

$$\begin{pmatrix} \Delta y_1 \\ \Delta y_2 \\ \Delta y_3 \\ \Delta k \end{pmatrix} = \frac{1}{4}\times\begin{pmatrix} 1 & -2 & -1 & 1 & -1 & 2 \\ -2 & 1 & -2 & 1 & 1 & 1 \\ -1 & -2 & 1 & -1 & 1 & 2 \\ 2 & 2 & 2 & 0 & 0 & -2 \end{pmatrix}\begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_6 \end{pmatrix}. \tag{65}$$

Applying covariance propagation law (29) to equation (65), the covariance propagation equation is obtained as follows:

$$\begin{pmatrix} \sigma^2_{\Delta y_1} & \sigma_{\Delta y_1 \Delta y_2} & \sigma_{\Delta y_1 \Delta y_3} & \sigma_{\Delta y_1 \Delta k} \\ \sigma_{\Delta y_2 \Delta y_1} & \sigma^2_{\Delta y_2} & \sigma_{\Delta y_2 \Delta y_3} & \sigma_{\Delta y_2 \Delta k} \\ \sigma_{\Delta y_3 \Delta y_1} & \sigma_{\Delta y_3 \Delta y_2} & \sigma^2_{\Delta y_3} & \sigma_{\Delta y_3 \Delta k} \\ \sigma_{\Delta k \Delta y_1} & \sigma_{\Delta k \Delta y_2} & \sigma_{\Delta k \Delta y_3} & \sigma^2_{\Delta k} \end{pmatrix} = \frac{1}{16} \times \begin{pmatrix} 1 & -2 & -1 & 1 & -1 & 2 \\ -2 & 1 & -2 & 1 & 1 & 1 \\ -1 & -2 & 1 & -1 & 1 & 2 \\ 2 & 2 & 2 & 0 & 0 & -2 \end{pmatrix} \mathbf{D}(\Delta \mathbf{X}) \begin{pmatrix} 1 & -2 & -1 & 2 \\ -2 & 1 & -2 & 2 \\ -1 & -2 & 1 & 2 \\ 1 & 1 & -1 & 0 \\ -1 & 1 & 1 & 0 \\ 2 & 1 & 2 & -2 \end{pmatrix}. \tag{66}$$

The acquisition process of the covariance matrix $\mathbf{D}(\Delta \mathbf{X})$ is as follows.

For the observed value x_i, its error Δx_i is composed of three parts: additive constant error K, multiplication constant error R, and uneven indexing error c_i. That is,

$$\Delta x_i = K + R \cdot x_i + c_i. \tag{67}$$

Its variance is

$$\sigma^2_{\Delta xi} = \sigma^2_K + x_i^2 \cdot \sigma^2_R + \sigma^2_c, \tag{68}$$

where σ_K, σ_R, and σ_c are obtained by consulting instrument instructions or the tolerance standard in instrument specification. Furthermore, according to formula (21), $\mathbf{D}(\Delta \mathbf{X})$ can be deduced as follows:

$$\begin{aligned} \mathbf{D}(\Delta \mathbf{X}) &= E \begin{pmatrix} \Delta x_1 \\ \Delta x_2 \\ \vdots \\ \Delta x_6 \end{pmatrix} \begin{pmatrix} \Delta x_1 & \Delta x_2 & \cdots & \Delta x_6 \end{pmatrix} \\ &= \begin{pmatrix} \sigma^2_K + x_1^2 \cdot \sigma^2_R + \sigma^2_c & \sigma^2_K + x_2 x_1 \cdot \sigma^2_R & \cdots & \sigma^2_K + x_6 x_1 \cdot \sigma^2_R \\ \sigma^2_K + x_1 x_2 \cdot \sigma^2_R & \sigma^2_K + x_2^2 \cdot \sigma^2_R + \sigma^2_c & \cdots & \sigma^2_K + x_6 x_2 \cdot \sigma^2_R \\ \vdots & \vdots & \vdots & \vdots \\ \sigma^2_K + x_1 x_6 \cdot \sigma^2_R & \sigma^2_K + x_2 x_6 \cdot \sigma^2_R & \cdots & \sigma^2_K + x_6^2 \cdot \sigma^2_R + \sigma^2_c \end{pmatrix}. \end{aligned} \tag{69}$$

Finally, the covariance matrix $\mathbf{D}(\Delta \mathbf{Y})$ is obtained by substituting equation (69) into equation (66), where $\sigma(\Delta y_j)$ is called as uncertainty.

Assuming that $\sigma_K = \pm 2\,\text{mm}$, $\sigma_R = \pm 1 \times 10^{-6}$, and $\sigma_c = \pm\, 1\,\text{mm}$, we can obtain

$$\begin{pmatrix} \sigma^2_{\Delta y_1} & \sigma_{\Delta y_1 \Delta y_2} & \sigma_{\Delta y_1 \Delta y_3} & \sigma_{\Delta y_1 \Delta k} \\ \sigma_{\Delta y_2 \Delta y_1} & \sigma^2_{\Delta y_2} & \sigma_{\Delta y_2 \Delta y_3} & \sigma_{\Delta y_2 \Delta k} \\ \sigma_{\Delta y_3 \Delta y_1} & \sigma_{\Delta y_3 \Delta y_2} & \sigma^2_{\Delta y_3} & \sigma_{\Delta y_3 \Delta k} \\ \sigma_{\Delta k \Delta y_1} & \sigma_{\Delta k \Delta y_2} & \sigma_{\Delta k \Delta y_3} & \sigma^2_{\Delta k} \end{pmatrix} = \begin{pmatrix} 0.75 & 0.01 & 0.26 & -0.50 \\ 0.01 & 0.78 & 0.05 & -0.50 \\ 0.26 & 0.05 & 0.85 & -0.50 \\ -0.50 & -0.50 & -0.50 & 5.00 \end{pmatrix} \left(\text{mm}^2\right). \tag{70}$$

Therefore, the uncertainties are

$$\begin{aligned} \sigma_{\Delta y_1} &\approx \pm 0.9\,\text{mm}, \\ \sigma_{\Delta y_2} &\approx \pm 0.9\,\text{mm}, \\ \sigma_{\Delta y_3} &\approx \pm 0.9\,\text{mm}, \\ \sigma_{\Delta k} &\approx \pm 2.2\,\text{mm}. \end{aligned} \tag{71}$$

It can be seen that $\sigma_{\Delta y_j} < \sigma_{\Delta x_i}$. Besides, it can be seen that the indication errors Δx_i not only lead to the dispersion of repeated observations $x_1 \sim x_6$, but also lead to their overall deviation. If we entangle in the influence characteristics (A/B classification) of indication error on repeated observations, it will not only be unable to express, but also will not help to solve the problem. Moreover, whether errors K, Rx_i, c_i, and Δx_i or error Δy_j are all deviations and have variances used to evaluate their probability intervals, the systematic error without variance does not exist.

Table 6: Conceptual logic difference between the two theories.

Traditional measurement theory	New conceptual measurement theory
The measured value is a random variable because it changes randomly in repeated measurements	The measured value is a numeric value and is a constant
The true value is a constant because it remains constant during repeated measurements	The true value is unknown and is a random variable that needs to be described by a probability range
The best measured value is given by random error analysis	The best measured value is given by analyzing the randomness of errors
With the true value as the reference center, submit the evaluation of the deviation and dispersion (reproducibility) of the measured value	With the measured value as the reference center, submit the evaluation value of the probability range that the true value deviates from the measured value
The errors are divided into systematic and random classifications	The errors are not divided into systematic and random classifications

8. Conclusions

In short, different from the traditional theory, the new theory follows rigorous mathematical concepts and regards both the observed value and measured value as constants and the error and true value as random variables so that the conceptual logic of new theory has changed in an overall way (shown in Table 6). First, the variance (standard deviation) or the uncertainty is the evaluation of probability interval of a single error (deviation), but not the dispersion of a measured value, and any regular error's size degree can be evaluated by them. Second, any error follows a random distribution, has variance which can be used to evaluate its size, and cannot be classified as the systematic error or the random error. Third, the error synthesis follows the algebraic rule, the variance synthesis follows the probability principle, and there is no need to use those old concepts such as systematic error, random error, precision, trueness, and accuracy. As a result, the revision of measurement textbooks and metrological terminology will become issues that need to be rediscussed in the future.

Data Availability

The data of the two cases in this paper are derived from historical data and are only used to describe the process of error processing. Other uses are invalid.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] JCGM 100:2008, *Guide to the Expression of Uncertainty in Measurement (GUM)*, BIPM, Sévres, France, Beijing, China, 2008.

[2] JJF1059-2012, *Evaluation and Expression of Uncertainty in Measurement*, China Quality Inspection Press, 2012.

[3] JCGM 200:2012, *International Vocabulary of Metrology—Basic and General Concepts and Associated Terms (VIM)*, BIPM, Sévres, France, 2012.

[4] JJF1001-2011, *General Terms in Metrology and Their Definitions*, China Quality Inspection Press, Beijing, China, 2011.

[5] GB/T14911-2008, *Basic Terms of Surveying and Mapping*, China Standard Press, Beijing, China, 2008.

[6] H. Huang, "A unified theory of measurement errors and uncertainties," *Measurement Science and Technology*, vol. 29, no. 12, Article ID 125003, 2018.

[7] X.-M. Ye, M. Ling, Q. Zhou, W.-N Wang, and X.-B. Xiao, "The new philosophical view about measurement error theory," *Acta Metrologica Sinica*, vol. 36, pp. 666–670, 2015.

[8] X.-M. Ye, X. Xiao, J. Shi, and M. Ling, "The new concepts of measurement error theory," *Measurement*, vol. 83, pp. 96–105, 2016.

[9] X. Ye, H. Liu, X. Xiao, and M. Ling, "The new concepts of measurement error's regularities and effect characteristics," *Measurement*, vol. 126, pp. 65–71, 2018.

[10] X.-M. Ye and S.-J. Ding, "Comparison of variance concepts interpreted by two measurement theories," *Journal of Nonlinear and Convex Analysis*, vol. 20, pp. 1307–1316, 2019.

[11] Y. Fei, *Error Theory and Data Processing*, Machinery Industry Press, Taichung, Taiwan, 2016.

[12] D. Ye, *Understanding, Evaluation and Application of Measurement Uncertainty*, China Metrology Publishing House, Beijing, China, 2007.

[13] School of Geodesy and Geomatics, Wuhan University, *Error Theory and Foundation of Surveying Adjustment*, Wuhan University Press, Wuhan, China, 2016.

[14] JJG703-2003, *Electro-Optical Distance Meter (EDM Instruments)*, China Metrology Press, Beijing, China, 2003.

[15] ISO 17123-4:2012, *Optics and Optical Instruments—Field Procedures for Testing Geodetic and Surveying Instruments—Part 4: Electro-Optical Distance Meters (EDM Measurements to Reflectors)*, ISO, Geneva, Switzerland, 2012.