ISSN 1330-3651 (Print), ISSN 1848-6339 (Online) https://doi.org/10.17559/TV-20220302123201
Original scientific paper

An Intelligent Screening Algorithm for Mining Key Dangerous Sources of Urban Ground Transport

Ruobing ZHANG, Xin ZHU, Tianshi WANG, Jing LI*, Xuejiao WANG

Abstract: With increasing bus capacity, operational intensity, etc., urban public transport emergencies are more and more characterized by heavy loads with high frequency. To build a collaborative public transport emergency command system (CPTECS) based on existing systems and datasets, bus emergency scenes and categories of sources of danger are defined. Emergency cases in Beijing are selected for analysis, designing new means of encoding and expanding the decision attributes of the rough set model. Cellular genetic algorithm (CGA) is used to screen key hazards highly correlated to existing information systems. By comparing with genetic algorithm (GA), it is found that CGA can better solve attribute reduction problems of multi-decision attribute rough set in stability, convergence quality, and algorithm efficiency. Based on the meteorological hazards screened out, a CPTECS is designed, enriching research in such territories. Research findings provide quantitative support for the design of CPTECS, and have certain practical significance.

Keywords: attribute reduction; cellular genetic algorithm; informatization collaboration; multi-decision attributes; public transport emergency; rough sets

1 INTRODUCTION

Public transport serves as an essential service of cities and has an indispensable role in the daily lives of the residents. Urban ground public transport has become an increasingly primary traveling method selection for its large capacity, flexibility, wide-coverage, high accessibility, low price, and low per capita fuel consumption. However, with the upgoing population, heavy loads, high frequency, and high intensity have become the prominent characteristics of the urban public transport environment, in which emergency occurs [1]. For example, in May 2002 in Chicago, USA, a bus collided with a truck carrying eight thousand gallons of gasoline, causing fire on the street and burning eight residential buildings; In June 2009 in Chengdu, China, 27 people were killed, and 74 were injured in a bus accident. Various public transport emergencies have occurred in major cities due to traffic accidents or natural disasters such as earthquakes, mudslides, and floods. Due to crowding, system vulnerability, and social sensitivity of public transport, careless handling of emergencies can cause severe traffic disruption, potentially triggering community dissatisfaction requiring intensive government action [2].

An emergency command system based on enterprise information and intelligence is currently being developed in Beijing. With information shared by entities including the Municipal Communications Commission Emergency Office, Municipal Meteorological Bureau, and Municipal Rail Transit Command Centre, significant human, material, and financial resources have been invested to establish a public transport information resource platform with 31 information sub-systems. Building an emergency command information collaboration system based on existing systems and data assets is still an urgent problem. However, not every emergency requires a highly coordinated response using information systems. To solve this problem requires analysis of the correlation between sources of danger and public transport information systems, screening of key bus emergency scenes, and development of a core business information-coordination mechanism for the construction of an urban ground transport emergency command system.

The remainder of this study is organized as follows: Section 2 introduces the related theory of rough sets, attribute reduction based on the GA, and the CGA. Section 3 provides a rough set attribute reduction model based on the CGA and applies it to an actual urban ground public transport emergency command. In Section 4, the performances of the GA and CGA are compared and analysed, and then the calculation results are as well analysed and thereafter an information collaboration system under critical emergency sources screened out by the model is built. In Section 5, the research results are summarized and further research directions are proposed.

2 BACKGROUND

2.1 Literature Review

Researchers have studied emergency management and emergency decision support systems [3], emergency rescue [4], and emergency management assessment [5]. The results provide a theoretical basis for subsequent research. However, existing research in the area of quantitative analysis and hazard screening is still insufficient.

Ground bus emergency command operation often involves many information systems of various importance, hence the difficulty to describe a complex knowledge system at the decision-making level uses qualitative methods [6], quantitative factor extraction methods [7, 8], or the traditional rough set model [9]. Therefore, public transport emergency event classification has questionable accuracy when based on sources of danger. However, the structure of "sources of danger-information system participation" is similar to that of the multi-decision attribute rough set model. Moreover, the algorithm's multi-decision attribute reflects the idea of system coordination, thus selected key emergency scenes meet the actual emergency coordination requirements. Unlike other methods that deal with uncertain problems, rough sets method does not need to be provided with a priori information beyond the processed dataset [10], and therefore is more suitable for emergency situations with incomplete or insufficient data.

This study takes public emergency events in Beijing from 2014 to 2016 and the corresponding emergency commands as the knowledge base, establishes a model based on the multi-decision attribute rough set theory, and abstracts the key emergency scenes into an attribute reduction problem.

Heuristic algorithms are usually used to solve the attribute reduction problems of rough sets. Common heuristic reduction methods include reduction algorithms based on the discrimination matrix [11] and information entropy [12]. However, heuristic algorithms are unsuitable for large-scale data calculations, where adaptive heuristic algorithms therefore apply. This is a kind of evolution algorithm that has been applied to various optimization problems [13]. The genetic algorithm uses group evolution and can be encoded to have an individual solution generated. To produce a new generation of individuals, the fitness function guides the search direction through selection, crossover, and mutation repeatedly, until the optimal solution is determined. Scholars have used GAs to reduce rough set attribute problems. [14] proposed an attribute reduction algorithm in the decision-theoretic rough set model (DTRSM) through region preservation that uses a GA for optimization. [15] extracted fault features based on the vibration signal generated by a gearbox, and used a GA-based rough set model to reduce attributes to obtain optimal fault characteristics and generate decision rules. The method can accurately and reliably detect the gearbox failure mode. The ant colony algorithm [16] and particle swarm algorithm [17] have also been applied to solve the attribute reduction problems of rough sets. Other meta heuristics algorithms (MA), such as monarch butterfly optimization algorithm [18], moth search algorithm [19], have also been proposed to solve global optimization problems recently. However, with complex problems, the above algorithms rapidly converge to local optima [20, 21]. Different kinds of implementations have been made to avoid such problems in recent MA study, such as Harris hawks optimization [22], slime mould algorithm [23], colony predation algorithm [24], hunger games search [25]. In another perspective, the idea of maintaining population diversity over a longer period of time for divided populations is proposed in order to avoid the previous problems [26], e.g., the cellular genetic algorithm. CGA used cellular automata to assign the population spatial structure, and the mutual learning mechanism can effectively protect excellent individuals, prevent resulting local extrema [27]. CGA has been widely used in dynamic optimization problems [28], and showed that it was superior to the general evolution algorithm applying to complex problems. Therefore, CGA is utilized in this study to reduce rough set attributes and screen key danger sources that have high correlation with the information system. Existing research combining rough sets and CGA is yet to be adequate.

Since the similarity between the construction of the decision table of a rough set and the encoding of a CGA, the genetic operation can be conveniently combined with the attribute reduction algorithm. A new encoding is used in this study, combining the attribute reduction problem of the multi-decision attribute rough set with a CGA, and a new reduction algorithm is formed. Through this method, key emergency scenes of public transportation are screened, key danger sources are matched with the information system, and the bus emergency command information coordination system is thereby explored.

2.2 Basic Theory of Rough Sets

At its core, the rough set associates a classification with knowledge and represents the classification of information by equivalent relations. Through knowledge reduction, that is, reducing redundant knowledge (attributes), the expression of the knowledge base becomes more concise without affecting the knowledge expression (Pawlak, 1992).

Definition 1: An information table can be described as (Pal and Skowron, 1999, Pawlak and Skowron, 2006) $S=\{U,C,V,D,f\}$, where $U=\{x_1,x_2,\cdots,x_n\}$ is the universe of discourse, $C=(c_1,c_2,\cdots,c_m)$ represents the condition attributes, $D=(d_1,d_2,\cdots,d_s)$ represents the decision attributes, $V=U_{(a\in C\cup D)}V_a$ is the range of the attributes, and $f=U\times(C\cup D)\rightarrow V$ is an information decision function. When $D\neq\varnothing$, the information table is a decision information table.

Definition 2: Indiscernible relation. For any attribute subset $B\subseteq(C\cup D)$,

$$IND(B)=\{(x,y)\in U^2 \mid f(x,a)=f(y,a),\forall a\in B\} \tag{1}$$

Definition 3: Equivalence class. The indiscernible relation can divide the universe of discourse $U/IND(B)$ (or U/B). Each divided area is an equivalence class, defined as

$$[x]_B=\{y\in U \mid (x,y)\in IND(B)\} \tag{2}$$

Definition 4: For any subset $X\subseteq U$ in any universe of discourse, the upper and lower approximation of *B* can be expressed as (Pawlak, 1982, 1992)

$$\underline{B}(X)=\{x\subseteq U\}=\{x\in U \mid [x]_B\subseteq X\} \tag{3}$$

$$\overline{B}(X)=\{x\subseteq U\}=\{x\in U \mid [x]_B\cap X\neq\varnothing\} \tag{4}$$

The upper approximations refer to the object set that can be determined to be classified in the class X in the *B*-based divided area. The lower approximations refer to the object set that may be classified in the class X in the *B*-based divided area.

For the attribute set $B\subseteq C$, the positive region, negative region (external region), and boundary region with respect to the decision set *D* are defined as

$$POS_B(D)=\cup_{X\in U/D}\underline{B}(X) \tag{5}$$

$$NEG_B(D)=U-\cup_{X\in U/D}\underline{B}(X) \tag{6}$$

$$BNG_B(D) = \cup_{X \in U/D} \overline{B}(X) - \cup_{X \in U/D} \underline{B}(X) \tag{7}$$

The positive region is the set of all regions that can be determined to divide into some *D*-based divided region in the *B*-based divided region, which reflects the classification capability of the attribute set *B* relative to *D*.

2.3 Cellular Genetic Algorithm Theory

First proposed the concept of the cellular genetic algorithm, the idea of which is that the evolution of an organism is not only related to the genetic material of the individual but also affected by the environment of its neighbours [29].

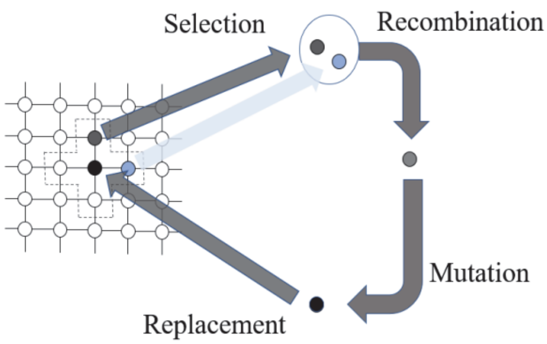

Figure 1 CGA theory

In the algorithm, the individuals comprising a group are assigned into a 2D grid (cell space) using cellular automata (CA). Unlike simple GAs, the selection and recombination operation of an individual is restrained in a limited neighbourhood by CGA, therefore decelerating the diffusion of genetic information throughout the population. Along with the evolution, the population is available to be more diverse through this method, consequently avoiding local convergence and responding to changes promptly.

3 SCREENING KEY SOURCES OF DANGER FOR PUBLIC TRANSPORT EMERGENCY BASED ON CGA

Urban ground public transport emergency command operation often involves the cooperation of many information systems of varying importance, and to describe such a system is difficult for the traditional rough set model of a single-decision attribute. In this section, the decision attributes of the rough set model based on the actual decision-making demands during a bus emergency are expanded, and an encoding condition attribute dependability and reduction algorithm based on CGA are designed.

3.1 Raw Data of Emergency Sources and Information System

207 actual emergency events and emergency command decisions in urban ground bus emergency command were collected from 2014 - 2016 to construct the urban underground public transport emergency command decision table shown in Tab. 3. Raw data can be seen in Tab. 1 and Tab. 2.

Table 1 Sources of danger and encoding

Encoding	Risk source
c_1	Road traffic accidents
c_2	Urban rail traffic accidents
c_3	Infrastructure and utility accidents
c_4	Mechanical accidents
c_5	Electrical accidents
c_6	Public security incidents
c_7	Mass incidents
c_8	Meteorological disasters

Table 2 Information system and encoding

Encoding	Information system
D_1	Beijing Bus Big Data Platform
D_2	Vehicle CAN Data Integrative Application Platform
D_3	Parcel Management System
D_4	Comprehensive Warranty Information Management Platform
D_5	Intelligent Operation and Dispatching System of Beijing Bus Group
D_6	Image Information Management System of Beijing Bus Group
D_7	BRT Intelligent Dispatching System
D_8	Beijing Bus Group Video Conference System
D_9	Group Emergency Management System
D_{10}	Business Real-time Data Forwarding Platform
D_{11}	Beijing Public Transport Information Resource Platform
D_{12}	Beijing Public Open Geographic Information Platform (GIS platform)
D_{13}	Operational Visualization Management System
D_{14}	Beijing Bus Group Inspection Management System
D_{15}	Line Network Center Business Management System
D_{16}	Beijing Bus Network Information Service System
D_{17}	Bus E-Link, APP
D_{18}	WeChat Microblog Platform
D_{19}	Ticket Management System
D_{20}	Human Resource Management System
D_{21}	Bus Security Information Management System
D_{22}	IT Operation and Maintenance Management Platform
D_{23}	Beijing Transportation Service Hotline System
D_{24}	Citizens' Opinion Management System
D_{25}	Group OA Office System
D_{26}	Vehicle Technology Management System
D_{27}	Safety Clothing Service Management System
D_{28}	Emergency Repair Integrated Information Management System
D_{29}	Publicity Business Management System
D_{30}	Security Hazard Investigation System
D_{31}	Letter and Visit System

3.2 Urban Ground Transport Emergency Command Decision Table

The bus emergency command knowledge system can be described as $S = \{U, C, V, D, f\}$, where $U = \{x_1, x_2, \cdots, x_n\}$ is a set of bus emergency events, $C = (c_1, c_2, \cdots, c_m)$ indicates the potential sources of danger that may trigger a bus emergency, $D = (d_1, d_2, \cdots, d_s)$ indicates the decision-making options involving related information systems in the bus emergency command, and $V = U_{(a \in C \cup D)} V_a$ is the region of attributes. In the sequence of condition attributes, "0" and "1" indicate respectively whether the occurrence of an emergency event does or does not involve a hazard. In the sequence of decision attributes, "0" and "1" respectively indicate whether this information system is or is not needed to handle a hazard. $f = U \times (C \cup D) \to V$ is an information decision function. The decision attributes are expanded based on the traditional rough set model, in order to construct a new rough set model with multi-decision attributes in this paper. $p_1, p_2, \cdots, p_s$ represent the relevance between a decision attribute and a condition attribute, which can be determined by calculating the grey relational grade, $p_j = \frac{1}{m}\sum_{i=1}^{m}\gamma_{ij}, j = 1, 2, \cdots, s$.

3.3 Dependability Between Key Sources of Danger and Information Systems of Public Transport Emergency

Definition 5: Let $K = (U, B)$ be an emergency command knowledge base, where B is a set of equivalence relations. $P, Q \subseteq U$, and the dependability between Q and P can be expressed as

$$k = \gamma_P(Q) = \frac{card(pos_P(Q))}{card(U)} \tag{8}$$

This paper modifies the dependability. In emergency command knowledge system $S = \{U, C, V, D, f\}$, $B \subseteq C$, $U = \{x_1, x_2, \cdots, x_n\}$. The dependability between the set of emergency event hazards and the set of information system attributes is expressed as

$$k(B, D) = \frac{1}{s}\cdot\sum_{i=1}^{s}\left[\begin{array}{l}\sum_{j=1}^{j}\max\left(\varphi_U(D_1)\cdots\varphi_U(D_j)\right)p_j \times \\ \times\dfrac{card\left(POS_{IND(B)}\left(IND\left(\{D_1,\cdots,D_j\}\right)\right)\right)}{card\left(POS_{IND(U)}\left(IND\left(\{D_1,\cdots,D_j\}\right)\right)\right)}\end{array}\right] \tag{9}$$

where p_j is the grey relational grade between the condition attribute and decision attribute,

$$p_j = \frac{1}{m}\sum_{i=1}^{m}\gamma_{ij}, 1 \le j \le s \tag{10}$$

$D_j(k) = \{d_i(1), d_i(2), \cdots, d_i(n)\}$ and $C_j(k) = \{c_i(1), c_i(2), \cdots, c_i(n)\}$ are sequences of decision and conditional attributes, respectively, in the emergency command knowledge system. Then the relational coefficient between $d_j(k)$ and $c_i(k)(j = 1, \cdots, d, i = 1, \cdots, m)$ is

$$\begin{aligned}\gamma_{ij}(k) = &\frac{\min_j \min_k \left|D_j(k) - C_i(k)\right|}{\left|D_j(k) - C_j(k)\right| + \rho \max_j \max_k \left|D_j(k) - C_i(k)\right|} + \\ &+ \frac{+\rho \max_j \max_k \left|D_j(k) - C_i(k)\right|}{\left|D_j(k) - C_j(k)\right| + \rho \max_j \max_k \left|D_j(k) - C_i(k)\right|}\end{aligned} \tag{11}$$

$\varphi_U(D_j)$ is the rough membership of U with respect to the condition attribute D_j. The concept of rough membership indicates the correlation between the related information system and the set of bus emergency events. The more frequent an information system occurs in the set of bus emergency events, the more important the information system is. The rough membership of the set of bus emergency events with respect to the decision attribute D_j is

$$\varphi_U(D_j) = \frac{card(U \cap D_j)}{card(U)} \tag{12}$$

It can be seen that Eq. (9) combines the rough membership between sets of attributes Eq. (12), the relational grade Eq. (11), and traditional dependability.

3.4 Attribute Reduction Process Based on CGA

In this study, condition attributes are sources of danger that may cause bus emergency events, and decision attributes are the information systems needed for the emergency command. In other words, screening the key danger sources with high emergency information cooperation is a process of reducing the condition attributes. Studies have shown that calculation results of GAs are unstable and easily fall into local optima when dealing with complex problems, where the CGA exceeds and is therefore widely used in dynamic optimization problems [28]. Thus, the attribute reduction problem described is treated as an operational optimization problem, where dependability is used as the decision target. CGA is used to reduce condition attributes by designing a new fitness function.

(1) Encoding

Related sources of danger $C = (c_1, c_2, \cdots, c_m)$ that may cause emergency events are encoded as binary strings (1, 0, 1, 0, 1, …) for individuals. If $c_i = 1$, the individual contains the ith condition attribute, vice versa, e.g., $C_1 = (0,1,1,0,1\cdots)$ indicates the factors c_2, c_3, c_5 , … are contained in chromosome C_1 .

(2) Fitness function

The dependability is used for the fitness function. The importance of the condition attribute rises as the dependability among a condition attribute and decision attribute grows, as well as the individual's fitness.

(3) Steps of the algorithm

Step 1: Produce an initial population.

In the l × l cell space, the l^2 condition attribute combination individuals are randomly generated and represented by x_{ij} . Among $i, j \in [1, l]$, m denotes the number of condition attributes. Then $x_{ij} = [c_{ij1}, \cdots, c_{ijk}, \cdots, c_{ijm}]$, and c_{ijk} randomly takes the value 0 or 1, where 0 means that the individual x_{ij} does not contain the attribute c_{ijk} , and 1 means the opposite.

Table 3 Information decision table of urban ground transport emergency command

Cases U	Condition Attribute C								Decision Attribute D				
	c_1	c_2	c_3	c_4	c_5	c_6	c_7	c_8	$d_1(p_1)$	$d_2(p_2)$	……	$d_{30}(p_{30})$	$d_{31}(p_{31})$
	Road traffic accidents	Urban rail traffic accidents	Infrastructure and utility accidents	Mechanical accidents	Electrical accidents	Public security incidents	Mass incidents	Meteorological disasters	Beijing Bus Big Data Platform	Vehicle CAN Data Integrative Application Platform	……	Security Hazard Investigation System	Letter and Visit System
x_1	0	0	0	0	1	0	0	0	1	1	……	1	0
x_2	0	0	0	0	1	0	0	0	1	1	……	1	0
x_3	0	0	0	0	1	0	0	0	0	1	……	1	0
x_4	0	0	0	0	0	0	1	0	0	0	……	0	1
x_5	0	0	0	0	0	1	0	0	0	0	……	0	0
x_6	0	0	0	0	1	0	0	0	0	0	……	1	0
x_7	0	0	0	0	0	1	0	0	0	0	……	0	0
x_8	0	0	0	0	0	1	0	0	0	0	……	0	0
……	……	……	……	……	……	……	……	……	……	……	……	……	……
x_{207}	0	1	0	0	0	0	0	0	0	0	……	0	0

Step 2: Calculate fitness.

Each cell calculates its own attribute condition dependability y_{ij} for the convenience of later comparison. This paper standardizes it so that the fitness of individual i, j is $fit_{ij} = y_{ij} - \min\limits_{i,j\in[1,l]} y_{ij}$.

Step 3: Selection.

Using the Moore-type neighbour structure $\left([i-1,i+1],[j-1,j+1]\right)$ expressed in Ω, each individual x_{ij} searches for the best individual in the neighbour $\left\{x_{st} | y_{st} = \max\limits_{i,j\in\Omega} y_{ij}\right\}$ as the learning object.

Step 4: Recombination.

The probability of recombination is set to p_c, and for each attribute c_{ijk} in x_{ij}, the probability p_c is interchanged with the attribute x_{st} in c_{stk}.

Step 5: Mutation.

The mutation probability is set to p_m, and for each attribute c_{ijk} in x_{ij}, the variation is generated by the probability p_m (1 is changed to 0, or 0 is changed to 1).

Step 6: Return to Step 2 until the fitness is no longer rising.

The pseudocode of evolution rules based on the CGA is also provided as **Algorithm 1** shown in Tab. 4.

Table 4 Pseudocode of evolution rules based on the CGA

Algorithm 1	CGA based on attribute reduction
1	While not termination Condition() do
2	For each cell i from 1 to population size
3	The attribute combination of cell i is coded by binary code
4	End
5	For each cell i from 1 to population size
6	Calculate the attribute dependability of cell i's attribute combination
7	Establish the fitness function based on dependability
8	Select the cell with the best fitness from the neighborhood i^*
9	Let p_c and p_m be the crossover and mutation probability, respectively
10	If $rand < p_c$
11	Implement the crossover operation between i and i^*
12	End
13	If $rand < p_m$
14	Implement the mutation operation in cell i
15	End
16	End
17	End

4 RESULTS AND ANALYSIS

This experiment was performed under the following development environment: MATLAB 2016Ra, CPU E5-1650 v2, graphics card GTX1060 6G Asu, memory Kingston 8G Recc, motherboard ASUS Z9PA, hard drive Samsung 850 EVO 250G SSD.

4.1 Performance Comparison of GA and CGA

This paper mainly compares the attribute reduction results of GA and CGA from the aspects of algorithm stability, convergence quality, and algorithm efficiency. The stability of the algorithm is measured by the number of occurrences of the global optimal solution in ten independent repeated reduction experiments. The quality of convergence is measured by the mean of the maximum fitness variance. The efficiency of the two algorithms is measured by the number of iterations required for convergence.

To compare the reduction results, this paper uses a GA to screen the emergency sources of danger, and the condition attribute combinations of the emergency sources of danger are encoded as binary strings that act as chromosomes (0 means that the attribute can be reduced, and 1 means it is a key attribute). In this paper, the same crossover probability and mutation probability are set for both algorithms. Both algorithms are used to reduce the bus emergency command decision table 10 times, i.e., 10 repetitions of the experiments are performed. After testing, it was found that both algorithms had converged after 50 iterations, so the maximum number of iterations selected in this paper is 50. The reduction process is shown in Tab. 5 and Tab. 6.

Table 5 Attribute reduction process based on GA

		Iteration 1	Iteration 2	Iteration 3
Reduction 1	Result	c_1, c_4, c_5, c_7, c_8	c_4, c_7, c_8	c_5, c_7, c_8
	Fitness	10.6894	11.08230	11.0372
Reduction 2	Result	c_2, c_3, c_5, c_6, c_8	c_2, c_5, c_8	c_5, c_8
	Fitness	10.6894	10.6563	10.4884
Reduction 3	Result	c_2, c_4, c_7, c_8	c_4, c_8	
	Fitness	10.6894	10.6225	
Reduction 4	Result	c_3, c_4, c_6, c_7, c_8	c_3, c_4, c_6, c_8	c_8
	Fitness	10.6894	11.0017	10.9862
Reduction 5	Result	c_6, c_8		
	Fitness	10.6894		
Reduction 6	Result	c_7, c_8	c_8	
	Fitness	10.6894	10.9769	
Reduction 7	Result	c_2, c_4, c_7, c_8	c_7, c_8	
	Fitness	10.6894	10.6225	
Reduction 8	Result	c_1, c_2, c_4, c_6, c_8	c_2, c_4, c_6, c_8	c_8
	Fitness	10.6894	10.6720	10.5748
Reduction 9	Result	c_4, c_6, c_8	c_6, c_8	
	Fitness	10.6894	10.9795	
Reduction 10	Result	c_1, c_4, c_5, c_6, c_8	c_5, c_6, c_8	c_6, c_8
	Fitness	10.6894	11.0350	10.9805

Table 6 Attribute reduction process based on CGA

		Iteration 1	Iteration 2	Iteration 3	Iteration 4
Reduction 1	Result	c_4, c_7, c_8	c_4, c_8	c_8	
	Fitness	10.6894	11.0372	11.0303	
Reduction 2	Result	c_1, c_3, c_5, c_6, c_8	c_1, c_5, c_8	c_1, c_8	c_8
	Fitness	10.6894	11.0193	11.0661	11.0172
Reduction 3	Result	c_1, c_2, c_4, c_5, c_8	c_1, c_2, c_4, c_8	c_4, c_8	c_8
	Fitness	10.6894	10.7240	10.6317	11.0240
Reduction 4	Result	$c_1, c_3, c_4, c_5, c_7, c_8$	c_1, c_4, c_5, c_8	c_4, c_8	c_8
	Fitness	10.6894	11.0678	11.0865	11.0240
Reduction 5	Result	c_1, c_2, c_7, c_8	c_1, c_8	c_8	
	Fitness	10.6894	10.6046	11.0172	
Reduction 6	Result	c_1, c_2, c_4, c_5, c_8	c_1, c_2, c_8	c_1, c_8	c_8
	Fitness	10.6894	10.7240	10.4677	11.0172
Reduction 7	Result	c_1, c_4, c_5, c_6, c_8	c_5, c_6, c_8	c_5, c_8	c_8
	Fitness	10.6894	11.0350	10.9805	11.0254
Reduction 8	Result	$c_2, c_3, c_4, c_5, c_7, c_8$	c_3, c_4, c_5, c_8	c_5, c_8	c_8
	Fitness	10.6894	10.7711	11.0565	11.0254
Reduction 9	Result	$c_1, c_2, c_3, c_4, c_5, c_8$	c_1, c_5, c_8	c_1, c_8	c_8
	Fitness	10.6894	11.0710	11.0661	11.0172
Reduction 10	Result	c_4, c_5, c_6, c_8	c_5, c_8	c_8	
	Fitness	10.6894	11.0260	11.0254	

Table 7 Types and numbers of occurrences of optimal solution

	Genetic algorithm					Cellular genetic algorithm
Optimal solution	c_4, c_8	c_5, c_8	c_6, c_8	c_7, c_8	c_8	c_8
Number of occurrences	1	1	3	2	3	10

Statistically analysing the attribute screening results of Tab. 5 and Tab. 6 to obtain Tab. 7, it can be seen that the GA has 10 independent repeated experiments with five types of solutions, while the CGA only has one type of solution, indicating that the GA easily falls into local optimal and is unstable. The CGA reached the global optimal solution in all 10 independent experiments, while the GA did so three times and fell into local optima seven times. This shows that the stability of the GA to obtain the global optimal solution is significantly worse than that of the CGA.

The degree of fitness can reflect the quality of an algorithm's convergence result when the optimal solution is determined. The greater the fitness the stronger the ability to find the optimal solution. The authors performed 50 iterations of the two algorithms and randomly selected a convergence result to draw Fig. 2. It is seen from the figure that the GA converges quickly and falls into a local optimum early. Although the CGA converges slowly, after the fifth iteration, it has significantly higher fitness than the GA. The maximum fitness of the CGA was 11.025, 0.537 higher than the GA's 10.488. This shows that the GA is premature and is trapped in a local optimum, hence it cannot obtain the global optimal solution, while the CGA has a stronger ability to find the global optimal solution.

In addition, the mean, standard deviation, and variance of the maximum fitness of the two algorithms were compared for 10 independent repeated experiments. The comparison results are shown in Tab. 8. Although they have the same maximum fitness, the average maximum fitness of the CGA is significantly higher than that of the GA, and it has a significantly smaller standard deviation and variance, indicating that its optimal solution is more stable.

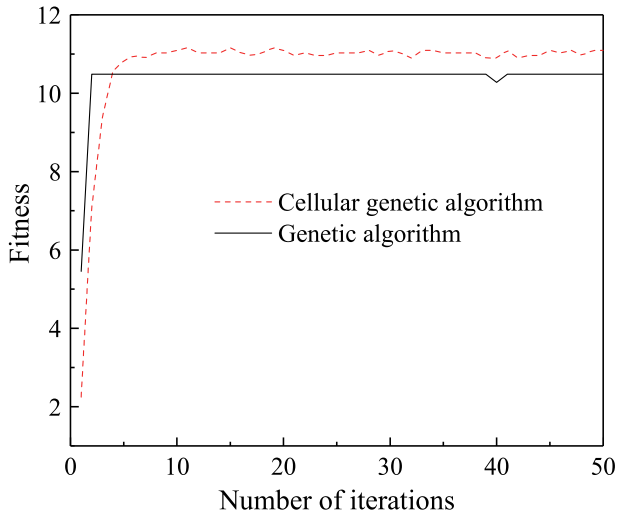

Figure 2 Fitness change of condition attribute reduction of GA and GCA

Table 8 Comparison of maximum fitness during evolution

Algorithm	Maximum fitness	Mean maximum fitness	Standard deviation of maximum fitness	Variance of maximum fitness
GA	11.0254	10.9310	0.1260	0.02
CGA	11.0254	11.0217	0.0037	1.3×10^{-5}

This study presents a boxplot of the optimal fitness of the optimal solution for 10 independent repeated experiments, as shown in Fig. 3. The authors can see that the CGA and GA have the same upper limit of the maximum fitness, but the difference between the upper and lower limits is quite different. The CGA has significantly higher stability than GA, confirming the comparison of Tab. 8.

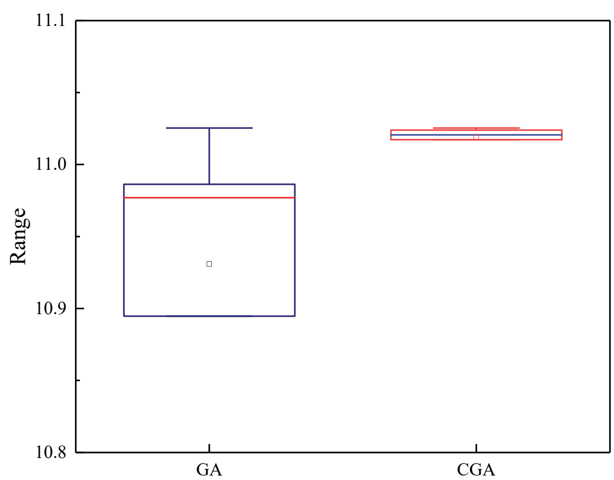

Figure 3 Boxplot of fitness of GA and CGA

The authors also compared the mean, standard deviation, and variance of fitness during 10 independent repeated experiments. A smaller fitness variance, as well as a smaller overall fitness mean square error, can reflect more stability of an algorithm in obtaining the optimal solution. To eliminate the influence of the initial value, the mean, standard deviation, and variance were calculated from the fitness values of the last 25 iterations of the optimal solution, and Tab. 9 shows the average of 10 independent repeated experiments. It can be seen that the average fitness of the CGA is 10.9834, which is 0.1711 more than that of the GA. However, the standard deviation and variance of the fitness of the CGA exceed those of the GA, indicating that the GA is prone to precocity and falls into local optima, so the fitness fluctuations in the evolution process are relatively small.

Table 9 Comparison of fitness for last 25 iterations in evolution process

Algorithm	Mean fitness	Standard deviation of fitness	Variance of fitness
GA	10.8123	0.0517	0.0033
CGA	10.9834	0.0657	0.0047

Finally, it can be seen from Fig. 2 that the convergence speed of the CGA is slightly lower than that of the GA. The GA converges in about three iterations, which is mainly because the CGA has a better global search ability.

Through the above comparison, it can be found that the GA converges faster than the CGA. However, the global search ability and stability of the CGA are higher. Therefore, the CGA performs better in solving attribute reduction problems of rough sets.

4.2 Key Sources of Danger and Related Information Systems

The principle of the attribute reduction algorithm is to examine the change of the decision table after removing an attribute, and then measure the importance of this attribute to the decision attribute.

Table 10 Key sources of risk and information systems involved

Key source of risk	Information systems involved
Meteorological disaster	1. Parcel Management System 2. Comprehensive Warranty Information Management Platform 3. Intelligent Operation and Dispatching System of Beijing Bus Group 4. Image Information 004Danagement System of Beijing Bus Group 5. Group Emergency Management System 6. Business Real-time Data Forwarding Platform 7. Beijing Public Transport Information Resource Platform 8. Beijing Public Open Geographic Information System (GIS platform) 9. Operation Visualization Management System 10. Line Network Center Business Management System 11. Beijing Bus Network Information Service System 12. Bus E-Link, APP 13. WeChat Microblog Platform 14. Emergency Repair Integrated Information Management System 15. Publicity Business Management System

The first attribute that is removed is least important. Based on the results of 10 independent experiments, it can be seen from Tab. 5 and Tab. 6 that the CGA maintains population diversity by introducing the cellular automaton, and can screen more accurate and reasonable key sources of danger than the GA. The final danger source based on the CGA is c_8, which represents meteorological disaster; thus, these are the most important key hazards in public ground transport emergencies in Beijing. Based on information from the Beijing Public Transport Group Co., Ltd., the bus emergency information systems involved in handling key sources of danger are shown in Tab. 10.

4.3 Meteorological Disaster Emergency Disposal Workflow and Information Coordination System Construction

When an emergency occurs, the effective implementation of the plan requires a high degree of coordination between information systems. Beijing city has invested abundant human, material, and financial resources to establish a public transportation information resource platform, which contains 31 sub-information systems. To better cope with emergencies, it is necessary to coordinate information systems.

Actual emergency data in Beijing is taken as an example. The emergency source of ground public transportation in Beijing is a meteorological disaster according to the model result. Based on investigations of the Beijing Public Transport Group, it is found that the emergency response to meteorological disasters involves external organizations such as the Transportation Commission, Meteorological Observatory, and news media. The disposition process mainly includes three links: receiving meteorological information, disposing, and reporting. The Beijing Public Transport Group formulates emergency plans based on forecast and early-warning information from the Meteorological Station and the news media, along with weather consultation from the Communications Commission. In the handling stage, when a meteorological disaster emerges, the emergency attendant immediately establishes a command or leading group, such as the flood control emergency command, snowy emergency work leading group, or high-temperature weather emergency team. After establishing the headquarters, the headquarters of the leading group consults with the Transportation Commission and the State-Owned Assets Supervision and Administration (SASAC) to implement the plan. Finally, in the reporting stage, the emergency management canter summarizes the emergency work for the Transportation Commission. The flowchart of emergency response work under meteorological disaster emergencies is shown in Fig. 4.

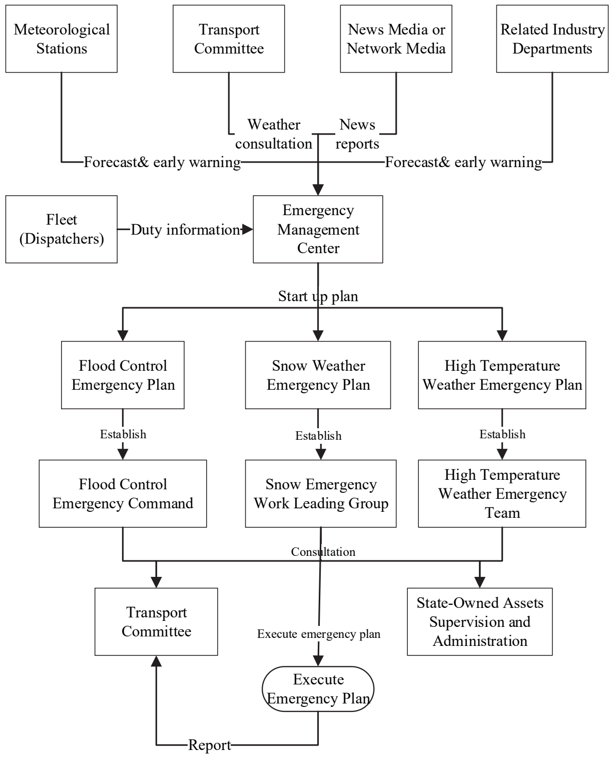

Figure 4 Flowchart of meteorological disaster emergency management

Based on the flowchart, information systems with high relevance to the meteorological disaster disposal process are selected to construct the collaboration system; these collaborative relationships are shown in Fig. 5.

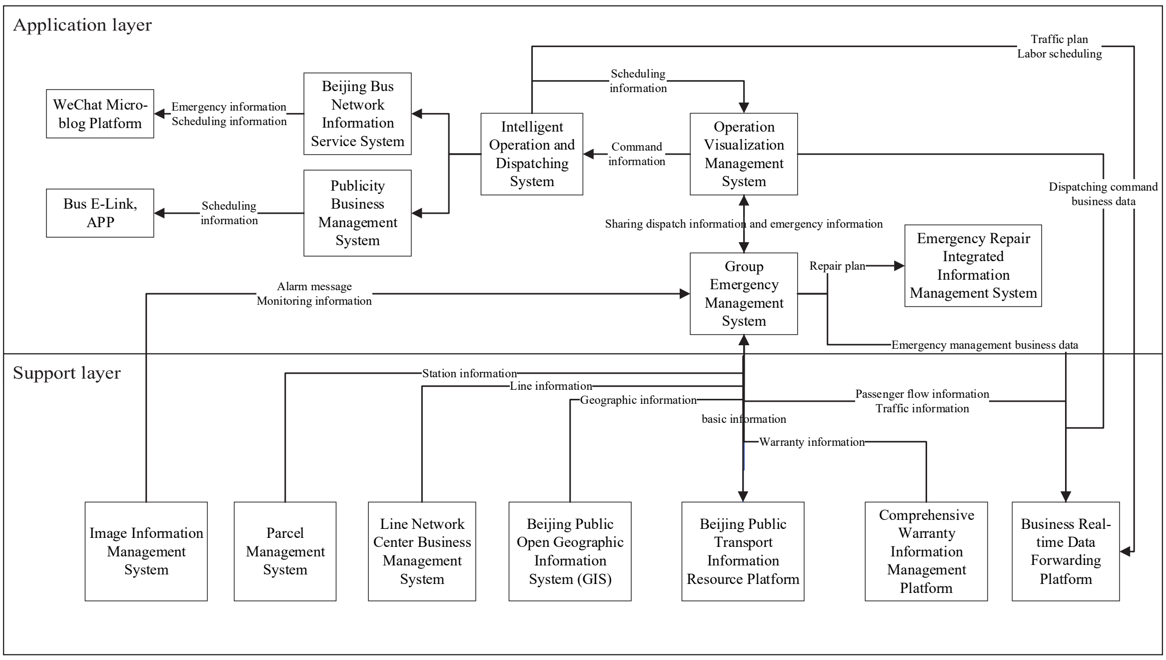

Figure 5 Collaborative relationship between information systems under meteorological disasters

The information system collaboration diagram is divided into support and application layers. The support layer mainly includes basic information such as about people, cars, lines, and stations, and a comprehensive data forwarding platform. The support layer aggregates and integrates basic information and converts it to a comprehensive information supply layer. The application layer consists of information systems directly related to emergency command, such as for emergency management, operational visualization, intelligent dispatching, line network service information, and propaganda business management. The system at the application layer plays a role in the emergency response process, and the information generated is only transferred between systems at the application layer, which reflects the characteristics of cooperative command.

At the support level, the Transit Information Resource Platform (SOA) receives basic information about people, cars, lines, and stations from geographic information, line networks, and parcel management systems, as well as business data from image systems, business forwarding platforms, and warranty systems, and forwards basic information to the emergency management system. The business real-time data forwarding platform manages real-time business data generated by all business systems, such as driving schedules and labour scheduling information of the intelligent dispatch system, dispatch command business data of the operation visual management system, and emergency management business data of the emergency management system. In addition, the business real-time data forwarding platform pushes information on passenger flow and road conditions to the emergency management system. The image information management system manages all alarm and video monitoring information and pushes it to the emergency management system. The warranty system pushes the warranty information of the vehicle to the emergency management system.

The emergency management, intelligent dispatching, and operation visualization systems form the core of the application layer. The intelligent dispatching system is used at the fleet level, and the operation visualization system is used by branches and group companies. In the event of a meteorological disaster, the emergency management center leads the emergency work, determines the emergency plan, issues emergency instructions to the operation visualization system, and simultaneously issues these to the intelligent dispatching system, which pushes the line adjustment information to the line network information service system and publicity business management system, which broadcast it to the public through WeChat, Weibo, and Bus Road. The intelligent dispatch and operation visual management systems share dispatch and command information, and the emergency management and operation visual management systems share command and emergency information. The emergency management system summarizes all emergency information and reports it to the Transportation Committee.

The established public transportation emergency management information coordination system has the following functions: (1) supervise and control each node of a meteorological disaster emergency, and realize the integration of dispatching command centre, sub-centre and emergency command vehicle-related hardware; (2) collect, transmit, and summarize information about existing GPS, GIS, video, and other management systems and auxiliary dispatching systems; (3) realize the collaboration of applications including visual command dispatch, video surveillance image display, and passenger information release; (4) improve the daily dispatch and command and comprehensive handling capabilities regarding meteorological disasters and emergencies to ensure the safe operation of urban ground public transportation.

5 CONCLUSION

Urban ground public transport emergencies have become increasingly common in major cities, making it imperative to build an emergency information coordination system. However, quantifying the relationship between sources of danger and information systems involved in emergency response decisions and screening out key sources of danger with high information coordination is vital to fully utilizing information systems and data resources and scientifically constructing an information collaboration system. In the urban ground bus emergency command, the structure of sources of danger, that is information system participation, is similar to that of decision support systems based on the multi-decision attribute rough set model, and the algorithm also reflects the idea of synergy. Moreover, the rough set theory requires no prior information beyond the dataset that must be processed and is suitable to situations of incomplete data in public emergencies. Therefore, this study abstracts the above problem as a 0-1 optimization problem and uses intelligent screening to identify key sources of danger based on collaborative information.

In total, 207 public transport emergency incidents and the corresponding emergency command in Beijing in recent years were selected as a knowledge base, and the problem of mining emergency key scenes as a multi-decision attribute reduction problem is abstracted. Considering the limitations of study methods based on the attribute reduction problem of rough sets, the authors designed a new coding method and decision attributes of the basic rough set model, combined the definitions of gray relevance, rough membership, and traditional dependency to form a new condition attribute dependence, and used a cellular genetic attribute reduction algorithm to extract key sources of danger, which are highly correlated with existing information systems. A comparison of the reduction performance of the CGA and GA showed that the former can obtain a more stable global optimal solution, while the latter easily and prematurely falls into local optima, indicating that the former performs better in solving attribute reduction problems of multi-decision attribute rough sets. Finally, based on the key hazard meteorological hazards screened out by the model, the authors designed a collaborative command system that patches the inadequacy in area of research on collaborative designs of information systems. At the same time, it provides ideas and references for the emergency response of the Beijing Municipal Public Transportation Group. It has a practical reference value to ensure the stable operation of public transportation, improve emergency response capacity, and ensure the stable operation of public transportation.

Acknowledgements

This work was supported by the Social Science Fund of Beijing (Grant No. 19ZDA05), the Youth Project of National Natural Science Foundation of China (Grant No. 72103019), and the Academic Research Projects of Beijing Union University (Grant No. ZK120202103).

6 REFERENCES

[1] Malik, F., Shah, M. A., & Khattak, H. A. (2018). Intelligent transport system: An important aspect of emergency management in smart cities. *2018 24th International Conference on Automation and Computing (ICAC)*, 1-6. https://doi.org/10.23919/IConAC.2018.8749062

[2] Pérez-Morales, A., Gomariz-Castillo, F., & Pardo-Zaragoza, P. (2019). Vulnerability of transport networks to multi-scenario flooding and optimum location of emergency management centers. *Water*, *11*(6), 1197. https://doi.org/10.3390/w11061197

[3] Abuova, A., Lakhno, V., Oshanova, N., Yagaliyeva, B., & Anosov, A. (2019, December). Conceptual model of the automated decision-making process in analysis of emergency situations on railway transport. *International Conference on Research and Practical Issues of Enterprise Information Systems*,153-162. https://doi.org/10.1007/978-3-030-37632-1_14

[4] Coles, D., Yu, D., Wilby, R. L. et al. (2017). Beyond "flood hotspots": Modelling emergency service accessibility during flooding in York, UK. *Journal of Hydrology*, *546*, 419-436. https://doi.org/10.1016/j.jhydrol.2016.12.013

[5] Poggi, V., Scaini, C., Moratto, L., Peressi, G., Comelli, P., Bragato, P. L., & Parolai, S. (2021). Rapid damage scenario assessment for earthquake emergency management. *Seismological Research Letters*. https://doi.org/10.1785/0220200245

[6] Nivolianitou, Z. & Synodinou, B. (2011). Towards emergency management of natural disasters and critical accidents: the Greek experience. *Journal of Environmental Management*, *92*(10), 2657-2665. https://doi.org/10.1016/j.jenvman.2011.06.003

[7] Huting, J., Reid, J., Nwoke, U. et al. (2016) Identifying factors that increase bus accident danger by using random forests and trip-level data. *Transportation Research Record: Journal of the Transportation Research Board*, (2539), 149-158. https://doi.org/10.3141/2539-17

[8] Huizenga, H. M., Zadelaar, J., & Jansen, B. R. (2021). Quantitative or qualitative development in decision making? *Journal of Experimental Child Psychology*, *210*, 105198. https://doi.org/10.1016/j.jecp.2021.105198

[9] Pawlak, Z., Grzymala-Busse, J., Slowinski, R. et al. (1982). Rough sets. *International Journal of Computer & Information Sciences*, *11*(5), 341-356. https://doi.org/10.1007/BF01001956

[10] Jing, Y., Li, T., Luo, C. et al. (2016). An incremental approach for attribute reduction based on knowledge granularity. *Knowledge-Based Systems*, *104*, 24-38. https://doi.org/10.1016/j.knosys.2016.04.007

[11] Qian, J., Miao, D. Q., Zhang, Z. H. et al. (2011) Hybrid approaches to attribute reduction based on indiscernibility and discernibility relation. *International Journal of Approximate Reasoning*, *52*(2), 212-230. https://doi.org/10.1016/j.ijar.2010.07.011

[12] Wang, F., Liang, J., & Dang, C. (2013). Attribute reduction for dynamic data sets. *Applied Soft Computing*, *13*(1), 676-689. https://doi.org/10.1016/j.asoc.2012.07.018

[13] Tan, F., Fu, X., Zhang, Y. et al. (2008) A genetic algorithm-based method for feature subset selection. *Soft Computing*, *12*(2), 111-120. https://doi.org/10.1007/s00500-007-0193-8

[14] Chebrolu, S. & Sanjeevi, S. G. (2011). Attribute reduction in decision-theoretic rough set models using genetic algorithm. *International Conference on Swarm, Evolutionary, and Memetic Computing*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-27172-4_38

[15] Huang, W., Kong, F., & Zhao, X. (2015). Spur bevel gearbox fault diagnosis using wavelet packet transform and rough set theory. *Journal of Intelligent Manufacturing*, 1-15. https://doi.org/10.1007/s10845-015-1174-x

[16] Jensen, R. & Shen, Q. (2004). Semantics-preserving dimensionality reduction: rough and fuzzy-rough-based approaches. *IEEE Transactions on Knowledge & Data Engineering*, (12), 1457-1471. https://doi.org/10.1109/TKDE.2004.96

[17] Wang, X., Yang, J., Teng, X. et al. (2007) Feature selection based on rough sets and particle swarm optimization. *Pattern Recognition Letters*, *28*(4), 459-471. https://doi.org/10.1016/j.patrec.2006.09.003

[18] Wang, G. G., Deb, S., & Cui, Z. (2019). Monarch butterfly optimization. *Neural computing and applications*, *31*(7), 1995-2014. https://doi.org/10.1007/s00521-015-1923-y

[19] Wang, G. G. (2018). Moth search algorithm: a bio-inspired metaheuristic algorithm for global optimization problems. *Memetic Computing*, *10*(2), 151-164. https://doi.org/10.1007/s12293-016-0212-3

[20] Katoch, S., Chauhan, S. S., & Kumar, V. (2021). A review on genetic algorithm: past, present, and future. *Multimedia Tools and Applications*, *80*(5), 8091-8126. https://doi.org/10.1007/s11042-020-10139-6

[21] Feng, Y., Deb, S., Wang, G. G., & Alavi, A. H. (2021). Monarch butterfly optimization: A comprehensive review. *Expert Systems with Applications*, *168*, 114418. https://doi.org/10.1016/j.eswa.2020.114418

[22] Heidari, A. A., Mirjalili, S., Faris, H., Aljarah, I., Mafarja, M., & Chen, H. (2019). Harris hawk's optimization: Algorithm and applications. *Future generation computer systems*, *97*, 849-872. https://doi.org/10.1016/j.future.2019.02.028

[23] Li, S., Chen, H., Wang, M., Heidari, A. A., & Mirjalili, S. (2020). Slime mould algorithm: A new method for stochastic optimization. *Future Generation Computer Systems*, *111*, 300-323. https://doi.org/10.1016/j.future.2020.03.055

[24] Tu, J., Chen, H., Wang, M., & Gandomi, A. H. (2021). The Colony Predation Algorithm. *Journal of Bionic Engineering*, *18*(3), 674-710. https://doi.org/10.1007/s42235-021-0050-y

[25] Yang, Y., Chen, H., Heidari, A. A., & Gandomi, A. H. (2021). Hunger games search: Visions, conception, implementation, deep analysis, perspectives, and towards performance shifts. *Expert Systems with Applications*, *177*, 114864. https://doi.org/10.1016/j.eswa.2021.114864

[26] Alba, E. & Tomassini, M. (2002). Parallelism and evolutionary algorithms. *IEEE transactions on evolutionary computation*, *6*(5), 443-462. https://doi.org/10.1109/TEVC.2002.800880

[27] Dorronsoro, B. (2013) Cellular genetic algorithms without additional parameters. *Journal of Supercomputing*, *63*(3), 816-835. https://doi.org/10.1007/s11227-012-0773-y

[28] Rezazadeh, H. & Khialimiab, A. (2017). A two-layer genetic algorithm for the design of reliable cellular manufacturing systems. *International Journal of Industrial Engineering Computations*, *8*(3), 315-332. https://doi.org/10.5267/j.ijiec.2017.1.001

[29] Whitley, L. D. (1993) Cellular genetic algorithms. *Proceedings of the 5th International Conference on Genetic Algorithms*. Morgan Kaufmann Publishers Inc. https://doi.org/10.5555/645513.657598

Contact information:

Ruobing ZHANG
School of Economics and Management, Beijing Jiaotong University,
Beijing 100044, People's Republic of China
E-mail: ruobing_zhang_hd@163.com

Xin ZHU
School of Management, Beijing Union University,
Beijing 100101, People's Republic of China
E-mail: 86369771@163.com

Tianshi WANG
School of Economics and Management, Beijing Jiaotong University,
Beijing 100044, People's Republic of China
E-mail: 20125642@bjtu.edu.cn

Jing LI
(Corresponding author)
School of Economics and Management, Beijing Jiaotong University,
Beijing 100044, People's Republic of China
E-mail: 18601107131@163.com

Xuejiao WANG
School of Management, Beijing Union University,
Beijing 100101, People's Republic of China
E-mail: 13453313593@163.com