U.P.B. Sci. Bull., Series C, Vol. 84, Iss. 1, 2022 ISSN 2286-3540

A DISTRIBUTED EVOLUTIONARY ALGORITHM FOR SOLVING THE GREEN SUPPLY CHAIN NETWORK DESIGN PROBLEM

Xinyuan LI[1], Dan WANG[2]*, Yanji JIANG[3], Maojun CAO[4]

This paper proposes an extended Multi-objective Mathematical Model for the Green Supply Chain (GSC) network design problem, taking green procurement, production, distribution, and transportation into consideration. The triple-bottom-line in GSC is defined and assigned indicators to evaluate potential partners for further operations. In addition, an efficient and flexible Self-Adaptive Distributed Evolutionary Algorithm (SADEA) is developed to undertake strategic, tactical, and operational decision making. The algorithm employs a hybrid tree coding structure with a self-adaptive operator selection pool and is paralleled and implemented in Spark to further improve its efficiency. Our results demonstrate that the algorithm can solve the GSC network design problem efficiently and provide high quality solutions compared to other algorithms for the multi-objective optimization. The study results provide an important reference to guide future research.

Keywords: green supply chain; evolutionary algorithm; distributed algorithm; selection pool; genetic algorithm; hybrid tree

1. Introduction

With a gradual recognition of the need to become sustainable, green supply chain management (GSCM) has received increasing interest from both academics and industry. Firms are challenged to balance business performance with environmental and social issues. To achieve this goal, companies have begun to incorporate sustainability concerns [1,2], increasing their demand for efficient and effective decision-making methods. Green design has gradually become the preponderant way for firms to sustainably develop.

A green supply chain's primary objective is to satisfy all customer demands via the most efficient use of resources, including inventory, logistics, and labor. Regarding the environmental aspect, carbon emission is commonly used as the indicator. However, studies tend to focus on only the emissions of parts of the supply chain process, such as production or transportation. This

[1] PhD eng., School of Business Administration, Liaoning Technical University, China
[2]* Prof., School of Business Administration, Liaoning Technical University, China, e-mail: wd-wangdan@163.com
[3] Prof., School of Software, Liaoning Technical University, China
[4] Prof., School of Business Administration, Liaoning Technical University, China

deficiency is primarily due to the complexity involved in modelling the various criteria. Optimization models become complex when combining social, environmental, and traditional business concerns. To account for sustainability in optimizing supply chain optimization, a large number of additional variables are required, rendering the solving of large-scale cases a serious challenge.

Due to the large number of indicators and partners in GSCM [3,4], the optimization of its network is difficult and time-consuming [5-9]. Thus, heuristic and metaheuristic algorithms, such as the Fuzzy Neural Network [10]; the Bees Algorithm [11]; the Ant Colony Optimization [12]; the Neural Network Algorithm [13]; and the Genetic Algorithm (GA) [14], have been widely used. However, four difficulties exist, rendering their use problematic. One, the algorithms remain time and memory-consuming in solving large-scale cases. Two, the diversity of the searching process usually decreases significantly after convergence to the local optimum, which restricts the ability of the algorithms to further explore the solution space. Three, the algorithms require an accurate parameter setting to obtain a good search ability, which can be difficult to obtain in a reasonable amount of time. Four, the algorithms have difficulties meeting the real-world requirements of scale, solution quality, and efficiency.

Thus, to further integrate sustainability into the supply chain network design, this study proposes a comprehensive Mixed Integer Linear Programming (MILP) model for a four-level green supply chain network design problem [15-17]. In addition, we propose a novel effective method combining a self-adaptive evolutionary algorithm, the optimum algorithm, and a distribution structure, based on a demand-oriented mechanism. We test its validity and reliability via a case study. Our results demonstrate that the proposed algorithm outperformed the other algorithms.

2. Literature Review

To develop a Green Supply Chain (GSC), an indicator system consisting of enterprise statistics should indicate corresponding changes and improvements of the existing supply chain [18]. The addition of carbon emissions is required to indicate the greenness of the GSC. Xie et al. [19] explored GSC performance indicators (i.e., energy consumption and profit), and Shao et al. [20] evaluated green performance using the energy usage rate and environmental pollution. However, these indicators do not completely describe the green performance. Coskun et al. [21] proposed a comprehensive indicator system for GSC performance for partner selection, introducing economic, environmental, and social performance, which is similar to the research of inventories [22] and research on facility allocation [23]. Roy et al. [24] proposed a two-level model containing one manufacturer and many retailers. Although attempts have been made to identify the indicators of green performance in a GSC [25,26], a holistic, sustainable framework emphasizing economic, environmental, and social dimensions does not yet exist.

Similar to traditional supply chains, the GSC can be designed as different models, with each designed for a different research direction. Aksen et al. [27] designed a discrete optimization model for solving the process of commodities trading in a green supply chain. Vencheh et al. [28] proposed a non-linear model for supplier selection in a green supply chain; later experiments demonstrated that the model was able to maintain stability during the solution process. Lious et al. [29] presented a model based on the green expectations of consumers. Benjaafar et al. [18] designed a model to balance the relationship between cost and carbon emissions. Roy et al. [30] proposed a two-level model with one manufacturer and many retailers. Wang et al. [31] proposed that a green supply chain required a comprehensive evaluation. As a result, the technologies of an environmental information system and a decision support system were synthesized in the research and a novel multi-objective dynamic programming model was proposed. Similarly, Lious et al. [29] designed a novel model for dealing with the Fuzzy information of decision-makers, while Oliveria et al. [32] and Sakiks et al. [33] built a multi-objective model that was consistent with the government's environmental policy. Kannan et al. [34] considered sustainable development using a modern production model. Xie et al. [35] proposed an integrated supply chain model that considered energy consumption and profit. Mirthedayatian et al. [36] evaluated double factors in a new DEA network model.

Many resolution algorithms have been proposed to design and optimize GSC networks [37]. For small-scale cases, Sarkis et al. [38] used the IBM ILOG CPLEX Optimization Studio (CPLEX) to solve the inventory problem. Wang et al. [31] and Boonsothonsati et al. [39] also used the CPLEX for the partner selection problem. Slimani et al. [40] proposed a game theory for a two-level supply chain with one retailer and one supplier. For medium-size cases, the Lagrangian relaxation, the Branch-and-bound approach, and the Bender decomposition have been utilized. However, these approaches cannot always provide feasible solutions. For large-scale cases, the solution space often depletes system memory during operation, resulting in the use of heuristic and metaheuristic methods. Kannan et al. [34] used the Genetic Algorithm to solve a recycling green supply chain model using a closed loop battery. The study also analyzed the closed loop supply chain using the Genetic Algorithm and the Particle Swarm Optimization Algorithm. Validi et al. [41] proposed a method based on the MOGA-II to solve the sustainable supply chain model. Roghanian et al. [42] optimized a network of reverse logistics using the Genetic Algorithm. Jiang et al. [43] solved the facility location problem using the Genetic Algorithm. Jamshidi et al. [44] used a hybrid memetic algorithm to improve the efficiency of the green supply chain network. Carlucci et al. [45] presented a non-dominated genetic algorithm for the multi-objective optimization of the supply chain. Dotoli et al. [46] proposed a multiple-step hierarchical algorithm for solving the supply chain network design problem with uncertainty.

Although it is allowable to significantly reduce the computational complexity of the search process by using heuristics and metaheuristics, the latter remains time/memory consuming in large-scale usage. Moreover, population-based algorithms require an accurate parameter setting to obtain a good search ability. Therefore, in this paper, we develop a self-adaptive distributed evolutionary algorithm to solve the green supply network design problem with high efficiency and solution quality, particularly in large-scale usage.

3. Mathematical model for GSC network design

Studies [15-17] helped develop and improve the criteria for evaluating the GSC network used in this paper. Table 1 displays the range of performance indicators and activity measures.

The hybrid method of the Analytic Hierarchy Process (AHP), as well as the Data Envelopment Analysis (DEA) method, were used to evaluate the weight of the selected indicators. The AHP is commonly used for organizing and analyzing complex decisions. It provides a rational framework for a needed decision by quantifying its criteria and alternative options, and for relating those elements to the overall goal. The DEA is a non-parametric method for performing frontier analysis. It uses linear programming to estimate the efficiency of multiple decision-making units and it is commonly used in production, management and economics.

The process was designed as follows: First, every index was compared after the hierarchical structure was established, and the judgment matrix was constructed according to the selected scale. Second, based on this judgment matrix, the single level ranking and the consistency check were implemented. After calculating the single ranking, we calculated the ranking weights of all the indicators in the same level, in order to obtain their relative importance at the highest level. The consistency of the general ranking was tested, and the weights were obtained. Then, the DEA model was established and transformed into an equivalent linear programming model. The model was then solved, and the optimal evaluation indicators and their weights were obtained. Finally, a linear combination as per Equation (1) was used to calculate the combined weights, where λ is the adjusting parameter.

$$\psi_i = \lambda\alpha_i + (1-\lambda)\beta_i \tag{1}$$

Table 2 displays the integrated indicators. Specifically, the cost-based indicators, such as worker rights and social compliance, were integrated directly into the corresponding partner's fixed costs. The probability-based indicators (e.g., example the accident rate) were transformed into a pessimistic value of lost costs and added into the corresponding partner's operational cost or production cost. The operational costs were calculated with indicators that included operational taxes, employee salaries, operational energy costs, and other costs of operation. Sustainability costs were calculated using the emission index, sewage discharge index, and pollution abatement.

Table 1

Indicators of each dimension considered in GSC

Bottom-line	Indicator	Type	Consolidation
Economic	Price	Positive real numbers	Costs of Production
	Quality	Positive natural numbers	Costs of Production
	Cost	Positive real numbers	Costs of Transportation
	Investment	Positive real numbers	Costs of Production
	Income	Positive real numbers	Costs of Operation
Environmental	Usage	Positive real numbers	Costs of Operation
	Energy cost	Positive real numbers	Costs of Production, Transportation and Operation
	Water cost	Positive real numbers	Water of Production, Operation
	Disposal	Positive real numbers	Costs of Operation
	Emission	Positive real numbers	Emission of Operation, Production and Transportation
	Reusability	Positive real numbers	Costs of Production
Social	Rights	Positive real numbers	Costs of Operation
	Risk	Positive real numbers	Costs of Operation
	Satisfaction	Positive natural numbers	Costs of Operation
	Compliance	Positive real numbers	Costs of Operation

Table 2

Integrated indicators in GSC

	Economic	Environment
Supplier	Raw material costs Transport costs	Transport emission
Manufacturer	Operational costs Product costs Transport costs	Product emission Operational emission Product water costs Operational water Transport emission
Distributor	Operational costs Stock costs Transport costs	Operational emission Operational water Transport emission

The proposed network of the green supply chain includes suppliers, manufacturers, distributors, and customers. The operation of the green supply chain was modelled as follows: The production sites process raw materials purchased from the suppliers and produce products that are then transported to a distribution center where the products are delivered to customers according to their orders. The objective was to obtain the balance between carbon emissions, water consumption, and total costs and find an inferior optimal solution in the feasible solutions. Thus, the proposed GSC network is a customer-order-driven process. Table 3 displays the indices of the supply chain partners, transportation modes, and energy types. Table 4 displays the decision variables in the GSC, and Table 5 displays the constants in the GSC.

Table 3

Variables in GSC

Variable	Number
Supplier	$i = \{1, 2, ..., I\}$
Manufacturer	$j = \{1, 2, ..., J\}$
Distributor	$k = \{1, 2, ..., K\}$
Customer	$l = \{1, 2, ..., L\}$
Transport	$s = \{1, 2, ..., S\}$
Energy	$e = \{1, 2, ..., E\}$

Table 4

Decision variable in GSC

Decision variables	Explanation
$s(m, f)t(d)_{j(k)}$	Binary variable that represents the current state of the manufacturer *j* (distributor *k*);
$as(t,d)t(d,c)_{i(j,k)j(k,l)s}$	Quantity of raw materials/product transported from one suppliers *i*(manufacturer *j*, distributor *k*) to manufacturers *j*(distributor *k*, client *l*) by transport s
$app(m)_{je}$	Numbers of products made by manufacturer *j* and energy *e*

Table 5

Constants in GSC

Constants	Explanation
D_l	Order of customer *l*
QS_i	Supply Capacity of supplier *i* to supply Raw material;
$QUT(D)_{j(k)}$	The capacity of manufacturer *j* (distributor *k*)
$CO(M.F)T(D)_{j(k)e}$	The fixed costs of manufacturer *j* (distributor *k*) by energy e
$IOT(D)_{j(k)}$	The State of manufacturer *j* (distributor *k*)
$CS(T,D)T(D,C)_{i(j,k)j(k,l)s}$	The cost of transporting a unit of raw material or product from supplier *i* (manufacturer *j*, distributor *k*) to manufacturer *j* (distributor *k*, client l) using transport mode *s*.
$CPP(M)_{je}$	The costs of producing a unit of product by manufacturer j and energy *e*
CSM_i	The costs of purchasing a unit of raw material from supplier *i*
$ES(T,D)T(D,C)_{i(j,k)j(k,l)s}$	Transport mode *s* from supplier *i* (manufacturer *j*, distributor *k*), transportation of a unit of raw materials or products to manufacturer *j*

	(distributor *k*, client *l*) that produces carbon emissions.
$EPP(M)_{je}$	The output of carbon dioxide produced by a unit of production by manufacturer *j* and energy *e*
$WPP(M)_{je}$	Manufacturer *j* uses energy *e* that produces the amount of water consumed by a unit of production
EPR_i	The amount of carbon dioxide produced by supplier *i* production of a unit of raw material
WPR_i	The amount of water consumed by supplier *i* for the production of a unit of raw material
$WO(M,F)T(D)_{j(k)e}$	Established (maintain, close) fixed water consumption of manufacturer *j* (distributor *k*) using energy *e*
$WS(T,D)I_{i(k)e}$	The water consumption indicator of supplier *i*(manufacturer *j*, distributor *k*)
$ES(T,D)I_{i(j,k)}$	The carbon emission indicator of supplier *i*(manufacturer *j*, distributor *k*)
$EO(M,F)T(D)_{j(k)e}$	Production (select, not select) fixed carbon dioxide emissions produced by manufacturer *j* (distributor *k*) using energy *e*
a,b	weights of the carbon emissions and weights of water consumption

Based on the above information, the objective function of cost, emissions, and water consumption were obtained using Equations (2) through (4):

$$obj1=\min(\sum_{s\in S}\sum_{k\in K}\sum_{l\in L}CDC_{kls}adc_{kls}+\sum_{s\in S}\sum_{j\in J}\sum_{k\in K}CTD_{jks}atd_{jks}+\sum_{s\in S}\sum_{i\in I}\sum_{j\in J}CST_{ijs}ast_{ijs}+\sum_{e\in E}\sum_{j\in J}CPP_{je}app_{je}+\sum_{i\in I}\sum_{j\in J}CSM_{i}ast_{ij}$$
$$+\sum_{e\in E}\sum_{j\in J}(CMT_{je}st_j+CFT_{je}ft_j)+\sum_{e\in E}\sum_{k\in K}(CMD_{ke}sd_k+CFD_{ke}fd_k)) \tag{2}$$

$$obj2=\min(\sum_{s\in S}\sum_{k\in K}\sum_{l\in L}EDC_{kls}adc_{kls}+\sum_{s\in S}\sum_{j\in J}\sum_{k\in K}ETD_{jks}atd_{jks}+\sum_{s\in S}\sum_{i\in I}\sum_{j\in J}EST_{ijs}ast_{ijs}+\sum_{e\in E}\sum_{j\in J}(EOT_{je}st_j+EMT_{je}st_jEFT_{je}ft_j)+$$
$$+\sum_{e\in E}\sum_{j\in J}(CMT_{je}st_j+CFT_{je}ft_j)+\sum_{e\in E}\sum_{k\in K}(EOD_{ke}sd_k+EMD_{ke}sd_k+EFD_{ke}fd_k)+\sum_{e\in E}\sum_{j\in J}EPP_{je}app_{je}+\sum_{i\in I}\sum_{j\in J}EPR_{i}ast_{ij}) \tag{3}$$

$$obj3=\min(\sum_{e\in E}\sum_{j\in J}(WOT_{je}st_j+WMT_{je}st_jWFT_{je}ft_j)+\sum_{e\in E}\sum_{k\in K}(WOD_{ke}sd_k+WMD_{ke}sd_k+WFD_{ke}fd_k)+\sum_{e\in E}\sum_{j\in J}WPP_{je}app_{je}$$
$$+\sum_{i\in I}\sum_{j\in J}WPR_{i}ast_{ij}) \tag{4}$$

The multiple objective programming model can be converted into a single objective programming model, as shown in Equation (5):

$$Obj=\min(obj1+a\times obj2+b\times obj3) \tag{5}$$

Subjected to the following constraints:

In this model, it was assumed that the orders were known and met, as shown in Equation (6):

$$\sum_{s\in S}\sum_{k\in K}adc_{kls}=D_l\,(l\in L) \tag{6}$$

The number of materials supplied from the suppliers must be less than the suppliers' capacity, as shown in Equation (7):

$$\sum_{s\in S}\sum_{j\in J} ast_{ijs} \le QS_i \left(i \in I\right) \tag{7}$$

The number of products from the suppliers to manufacturers must equal the number of products from manufacturers to distributors, as shown in Equation (8):

$$\sum_{i\in I}\sum_{j\in J} ast_{ijs} = \sum_{j\in J}\sum_{k\in K} atd_{jks} \tag{8}$$

The number of products from the suppliers to manufacturers cannot exceed the number of the products by manufacturers, as shown in Equation (9):

$$\sum_{s\in S}\sum_{i\in I} ast_{ijs} = \sum_{e\in E} app_{je} \tag{9}$$

The amount of carbon emissions by the suppliers cannot exceed the carbon emission index, as shown in Equation (10):

$$\sum_{i\in I} ESI_i \ge \sum_{s\in S}\sum_{j\in J} EPR_i ast_{ijs} \tag{10}$$

The amount of carbon emissions by the manufacturers cannot exceed the carbon emission index, as shown in Equation (11):

$$\sum_{j\in J} ETI_i \ge \sum_{s\in S}\sum_{j\in J} (EOT_j st_{jks} + EMT_j st_{jks} + EFT_j\, ft_{jks}) + \sum_{j\in J}\sum_{s\in S} EPP_j \times \sum_{i\in I}\sum_{j\in J} ast_{ijs} \tag{11}$$

The amount of carbon emission by the distributors cannot exceed the carbon emission index, as shown in Equation (12):

$$\sum_{j\in J} EDI_i \ge \sum_{l\in L}\sum_{k\in K} (EOD_k sd_{kl} + EMD_k sd_{kl} + EFD_k\, fd_{kl}) \tag{12}$$

The amount of water consumption by the suppliers cannot exceed the water consumption index, as shown in Equation (13):

$$\sum_{i\in I} WSI_i \ge \sum_{s\in S}\sum_{j\in J} WPR_i ast_{ijs} \tag{13}$$

The amount of water consumption by the manufacturers cannot exceed the water consumption index, as shown in Equation (14):

$$\sum_{j\in J} WTI_i \ge \sum_{s\in S}\sum_{j\in J} (WOT_j st_{jks} + WMT_j st_{jks} + WFT_j\, ft_{jks}) + \sum_{j\in J}\sum_{s\in S} WPP_j \times \sum_{i\in I}\sum_{j\in J} ast_{ijs} \tag{14}$$

The amount of water consumption by the distributors cannot exceed the water consumption index, as shown in Equation (15):

$$\sum_{k\in K} WDI_i \ge \sum_{l\in L}\sum_{k\in K} (WMD_k sd_{kl} + WFD_k\, fd_{kl}) \tag{15}$$

The open/close status of the partners are shown in Equations (16) to (28), where M is a very large value.

$$M \times st_j - \sum_{e\in E} app_{je} \ge 0\left(j \in J\right) \tag{16}$$

$$M \times st_j - \sum_{i\in I}\sum_{j\in J} ast_{ijs} \ge 0\left(j \in J\right) \tag{17}$$

$$M \times st_j - \sum_{s \in S} \sum_{k \in K} atd_{jks} \geq 0\left(j \in J \right) \tag{18}$$

$$M \times sd_k - \sum_{s \in S} \sum_{k \in K} atd_{jks} \geq 0\left(k \in K \right) \tag{19}$$

$$M \times sd_k - \sum_{s \in S} \sum_{l \in L} adc_{kls} \geq 0\left(k \in K \right) \tag{20}$$

$$st_j + \sum_{e \in E} ft_{je} \leq 1(j \in J) \tag{21}$$

$$st_j - \sum_{e \in E} ot_{je} \geq 0(j \in J) \tag{22}$$

$$sd_k + \sum_{e \in E} fd_{ke} \leq 1(k \in K) \tag{23}$$

$$sd_k - \sum_{e \in E} od_{ke} \geq 0(k \in K) \tag{24}$$

$$st_j \leq IOT_j + \sum_{e \in E} ot_{je} \leq 1(j \in J) \tag{25}$$

$$sd_k \leq IOD_k + \sum_{e \in E} od_{ke} \leq 1(k \in K) \tag{26}$$

$$1 - st_j \leq (1 - IOT_j) + \sum_{e \in E} ft_{je} \leq 1(j \in J) \tag{27}$$

$$1 - sd_k \leq (1 - IOD_k) + \sum_{e \in E} fd_{ke} \leq 1(k \in K) \tag{28}$$

4. Self-adaptive distributed evolutionary algorithm

In each iteration of the evolution process, the algorithm must adaptively select an operator pair from the pool to modify the current population of solutions. The operator pool contains several pairs of different crossover and mutation operators. The self-adaptive selection of the operator pairs is determined by the update of the global best solution, which is designed to guide the behavior of the searching process. The objective of this mechanism is promoting the algorithm to find a faster and more efficient solution.

4.1 Encoding

A hierarchy of the tree structure is used to demonstrate the conformation model of the constraint network. The chromosome is encoded as a multiway tree. Each tree is expressed using sequential numbers where each number indicates a partner in the GSC. Fig. 1 shows an example of the chromosome structure.

4.2 Initial solution generation

The initial solution of the proposed algorithm is obtained by sorting each partner's preference evaluation value, which is calculated by weights using the AHP and the DEA. The probability of each partner selected is calculated using Equation (29), where p_i represents the fitness of the potential partner.

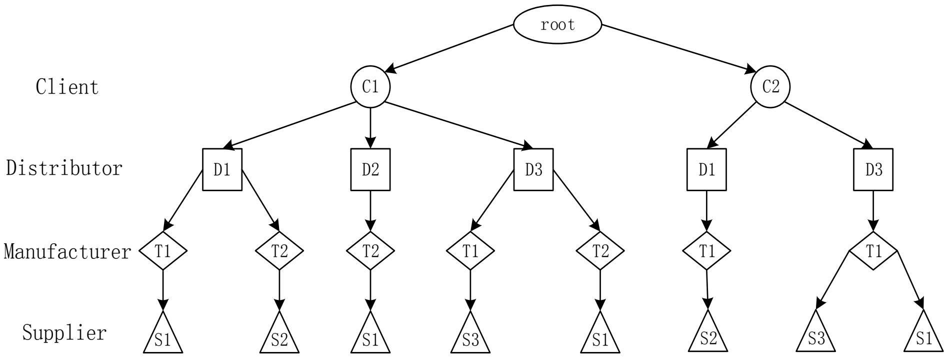


Fig 1. Tree structure of the chromosome

$$q_i = \sum_{j=0}^{N}((1-p_i^{-1} / \sum_{j=0}^{N}1-p_j^{-1})) \tag{29}$$

The cumulative probability of each supply chain partner, is calculated using p_k , as shown in Equation (30):

$$p_k = \sum_{j=1}^{K} q_j,\ j = 1, 2..., k \tag{30}$$

The probability of a random generation of fitness is within the range of $\left(p_{k-1}, p_k\right)$. Then, the same level kth partner will be selected as a member of the supply chain.

4.3 Evolutionary operations

In the crossover operation, it is assumed that $T_1(r_1, BT_1)$ and $T_2(r_2, BT_2)$ are the progenitors of the $T_1^{'}$ and $T_2^{'}$, respectively, where r_1 and r_2 are roots, BT_1 is the subtree of r_1, and BT_2 is the subtree of r_2. The crossover operation was performed, as indicated in Equation (31):

$$\begin{cases} T_1^{'} = \alpha T_1(r_1, BT_1) + (1-\alpha)T_2(r_2, BT_2) \\ T_2^{'} = (1-\alpha)T_1(r_1, BT_1) + \alpha T_2(r_2, BT_2) \end{cases} \tag{31}$$

where α is the partial probability, and p_c represents the crossover rate. In the mutation operation, when the $random > p_m$,

$$T_1^{'} = T_1(r_1, BT_1) \quad i = 1, 2, ..., n \tag{32}$$

where r_i represents all of the root records that meet the selection criteria; and p_m is the mutation rate.

Based on the structure of the chromosome, ILOG CPLEX is then applied to adjust the quantity of products transported between each of the two partners. Adopted from the structure introduced in Meignan et al. [51], the genetic operator classifier selection mechanism in the GSC was designed to be based on the form of ⟨condition, operation⟩, where the condition indicates that the current situation occurs, and the operator corresponds to the pair of crossover and mutation operations. Let C be the set of conditions and O be the crossover and mutation operator pairs. For a parameter condition, c_i, a genetic manipulation, o_i, is selected for each pair. The selection of an operator pair is performed by a counter parameter, which represents the distance number of iterations since the last time the local best individual was improved. The different operation pairs use the same crossover and mutation mechanism introduced above; however, their parameters are set to be different from each other to obtain different intensification and diversification search abilities.

4.4 Distributed structure

In the distributed structure, the population is divided into four sub-populations. When a sub-population searches around a local optimum, it may be discovered, while the other sub-populations continue to search for new local optima. The process is repeated if more local optima are found. The interaction among the agents of the distributed structure can maintain the diversity after a convergence. Indeed, exchanging information between cooperative meta-heuristics will alter their behavior in terms of searching in the landscape associated with the problem. Both better convergence and improvement in the quality of solutions may occur. Furthermore, by proceeding with the calculation on multiple workstations, the distributed structure can improve the computational speed of the algorithm. Fig. 2 illustrates the structure of this method.

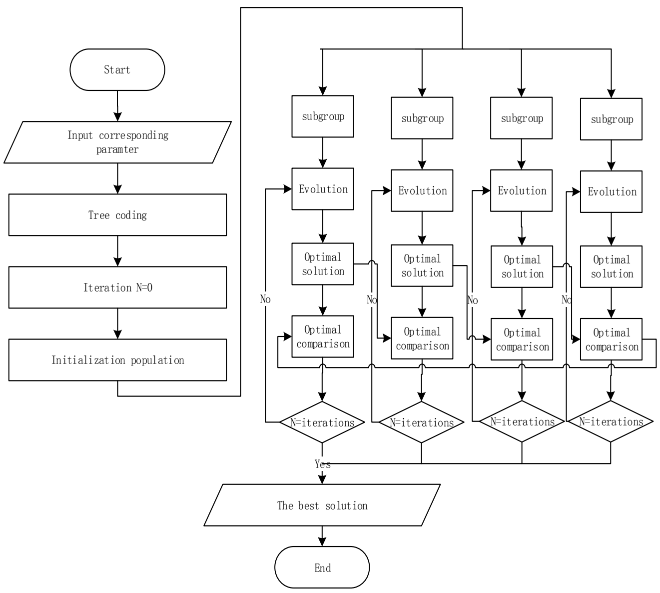


Fig. 2. The structure of Distributed Evolutionary Algorithm

5. Results and analysis

The experiments were conducted on a set of cases with different scales. The partners' maximum capacities of each case were generated according to real-world cases. The parameter setting and simulation configuration for the investigated approach were specified as follows: Population: 100, Generations: 500; $p_c = 0.6$, $p_m = 0.01$; $c_o = 10$; $c_1 = 30$; $c_2 = 60$; $c_3 = 95$; $o_0 : m_0 = 5.0\%$ of the individual's subtrees; $o_1 : m_1 = 10.0\%$ of the individual's subtrees; $o_2 : m_2 = 15.0\%$ of the individual's subtrees; $o_3 : m_3 = 20.0\%$ of the individual's subtrees.

Six sets of experiments were performed to evaluate the efficiency and effectiveness of this approach, and the number of vertexes in these sets increased gradually. For each set, five cases with respect to the real data sets of an electronic device manufacturing process were generated. Each case was run on the Genetic Algorithm (GA), the Cplex (Opt), the Self-Adaptive Evolutionary Algorithm without distributed structure (SAEA), and the SADEA; their average objective function (Avg), standard deviations (Std), and average run times (Time) were recorded for ten runs.

Table 6 displays a solution provided by the SADEA. It includes product distribution schemes among partners, and the results calculated using Equations (2) to (4). This solution corresponds to the data of a manufacturer of electronic goods that plans its GSC network to better meet customer requirements. The manufacturer has three customer regions, manufacturing factories, and distribution facilities, and three suppliers in the region.

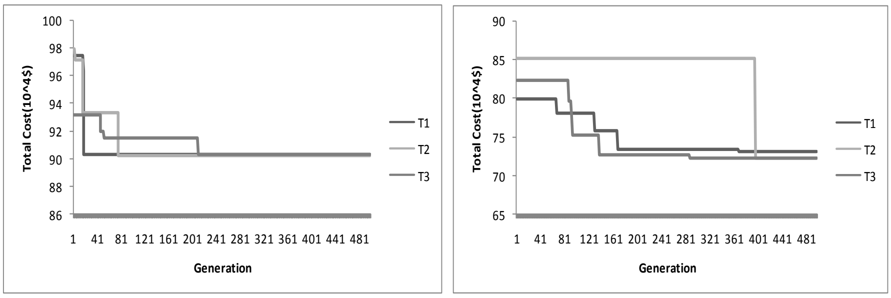
Table 6

Solution provided by SADEA

Partner	T3	T1	T2	D1	D2	C1	C2	C3	Total
S3	100								
S2		150							
S1		50	150						
T3				100					
T1				50	150				
T2					150				
D1								150	
D2						100	150	50	
Order						100	150	200	
Cost						176000	189000	159000	524000
Emission						643245	746060	876180	2265485
Water						7300	6900	8100	22300
Total						311949	345112	342336	999397

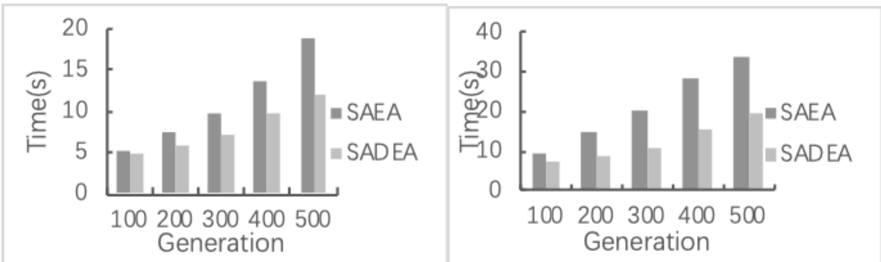
Fig. 3 shows the relationship between the total costs and the different parameter settings of the selection pool at the three scales. The parameters were set as follows: $T_1(C_1 = 10; C_2 = 25; C_3 = 50; C_4 = 75)$, $T_2(C_1 = 15; C_2 = 30; C_3 = 60; C_4 = 90)$, $T_3(C_1 = 20; C_2 = 40; C_3 = 80; C_4 = 100)$. It can be seen that the parameters affect the solution quality, and T_1 was more efficient than the other parameters.

Fig. 4 compares the SAEA and SADEA on the 2-node Spark cluster. It is obvious that SADEA's run time was only half that of the SAEA. By observing Fig. 5, we can find that the computing time of each cluster was slightly lower than the time of the stand-alone for small-scale cases. However, with the gradual increase in the case size, the advantages of the distributed computing began to appear.



(1) 30/30/30 sets (2) 80/80/80/80 sets

Fig. 3. The optimum of different instance sets



(1) 30/30/30 sets (2) 80/80/80/80 sets

Fig. 4. The Run time of different instance sets

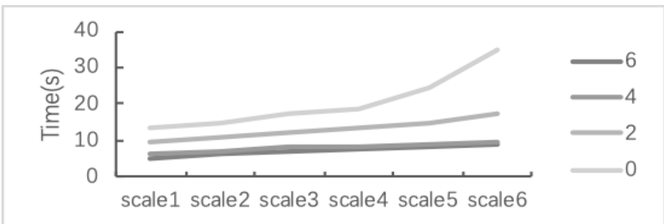


Fig. 5. Comparison run time in different clusters

Through the clustered calculation, the computing time of the algorithm can be significantly reduced, and the efficiency of the 6-node cluster provided the shortest computing time. For the last set of cases, the cluster with 4-nodes had a two-fold reduction in the computing time compared with the stand-alone. As shown in Table 7, SADEA was more efficient than the other algorithms. It can solve large-scale cases, which the CPLEX cannot. Furthermore, when the scale of the case increased, the computing time became significantly less than the other algorithms.

Table 7-1

Comparison of test results of each algorithm

Partner (S/P/D/C)	**Instances**	**Opt**	**GA**			**SAEA**			**SADEA**		
			Avg	Std	Time	Avg	Std	Time	Avg	Std	Time
5/5/5/5	1	910141.7	916580	4757.317	9	914105.3	3139.809	12	910382.6	9362.654	5
	2	891034.5	911355.2	4420.055	11	902294.3	6063.129	12	893727.4	3822.533	6
	3	909873.2	924256.9	4851.49	10	919710.6	7405.302	14	923846.7	22918.424	7
	4	910286	922222	5112.823	11	915736.7	3800.453	10	914638.5	21932.532	7
10/10/10/10	1	879573.1	912500.5	12710.31	16	900333.2	16925.2	15	894729.4	31737.432	9
	2	930928.6	945242.3	7045.947	15	940937.2	6654.129	14	940184.5	38274.435	11
	3	941083.2	953618.5	4620.07	14	949277.4	5220.199	16	947294.8	32948.258	10
	4	892537.1	923510.9	4737.444	14	904894	5422.064	16	923448.9	22473.854	10
30/10/10/10	1	832014.8	851385.2	13341.02	15	847070.5	8354.608	16	843929.4	38264.895	10
	2	969837.3	978736.1	6596.852	16	974194.8	3916.702	17	970392.1	40285.938	10
	3	937832.9	951073.9	5314.941	16	947710.1	4029.417	16	950284	28372.927	11
	4	926739.5	945947.2	9214.81	17	937825	5641.173	18	929948.9	24958.294	10
30/30/30/30	1	---	974643	7723.046	20	962753.5	3868.29	20	968482.3	21938.028	12
	2	---	943504.1	8060.757	20	933361.9	6667.377	21	924938.7	35288.384	11
	3	---	868483.6	14883.89	18	843349.8	6212.682	20	838492	38472.384	10
	4	---	946771.7	12728.5	18	934366.1	10503.52	23	938482.9	24858.937	10
50/50/50/50	1	---	946884	7803.034	21	939033.2	5110.345	26	923948.6	30991.969	16
	2	---	949144.9	7641.312	23	931613	2812.201	28	940028.4	32127.092	16
	3	---	955298.7	13680.2	22	932986.9	14367.71	25	929384.3	34153.374	15
	4	---	965989.8	17424.86	25	947854.8	7397.529	24	940294.8	32034.867	16
80/80/80/80	1	---	956664.5	26400.85	28	955067.7	8660.9	36	940284.5	23928.839	19
	2	---	929341.5	15594.61	29	907040	8881.099	38	899937.2	32484.982	20
	3	---	939958.3	14859.02	30	912081.8	12782.01	34	913828.6	29372.028	22
	4	---	957822.5	24108.96	29	914959.7	4146.388	38	903829.3	23947.493	21

Table 7-2

Comparison of test results of each algorithm (improvement)

Partner (S/P/D/C)	**Instances**	**Opt**	**Improve Opt(-%)**	**Improve GA(%)**	**Improve SAGA(%)**
			Opt(-%)	GA(%)	SAEA(%)
5/5/5/5	1	910141.7	0.02646841	0.67614393	0.40725067

	2	891034.5	0.30222174	1.93424035	0.9494574
	3	909873.2	1.53576344	0.0443816	-0.4497176
	4	910286	0.47814643	0.82230743	0.1199253
10/10/10/10	1	879573.1	1.72314274	1.94751674	0.62241401
	2	930928.6	1.75694462	0.53507974	-0.6745721
	3	941083.2	0.66004791	0.66312682	0.2088536
	4	892537.1	3.51938312	0.00671351	-0.1057605
30/10/10/10	1	832014.8	3.83582119	0.87572582	-1.9902594
	2	969837.3	0.98519618	0.85252807	-0.533497
	3	937832.9	2.18067632	0.08305348	-0.1157315
	4	926739.5	0.34631091	1.69124662	0.83982619
30/30/30/30	1	---	---	0.63209811	-0.5950433
	2	---	---	1.9677074	0.90245809
	3	---	---	3.45332946	0.57601247
	4	---	---	0.87548033	-0.4405982
50/50/50/50	1	---	---	2.42219744	1.60639688
	2	---	---	0.96049613	-0.903315
	3	---	---	2.71270127	0.38613618
	4	---	---	2.65996597	0.79759052
80/80/80/80	1	---	---	1.712199	1.54786933
	2	---	---	3.163993	0.78307462
	3	---	---	2.77987864	-0.1915179
	4	---	---	5.63707785	1.21649074

6. Conclusions

In this study, an integrated mathematical model and a new Self-Adaptive Distributed Evolutionary Algorithm were proposed and applied on a green supply chain network design problem. Related indicators were integrated into three different criteria that were used to evaluate the partners. A hybrid tree coding evolutionary algorithm was applied with a self-adaptive operator selection pool to facilitate the parameter setting and to improve the search ability of the algorithm. Finally, the approach was paralleled and implemented in Spark to further improve its efficiency and solution quality. Experiments were conducted to evaluate the performance of this approach, and the results showed that this approach can provide high quality solutions efficiently in large-scale cases.

Future research should consider more indicators in a green supply chain and extend this model by considering the uncertainty in customer demand and facility capacity. This will enhance its applicability in real-life scenarios. Closed-loop SCNs should also be included in future research to consider used/returned products flows.

Acknowledgments

This study was supported by the Foundation of Liaoning Province Education Administration [Grant number LJ2020FWL001].

REFERENCES

[1] *Weng, H.H.R., Chen, J.S., Chen, P.C.* Effects of green innovation on environmental and corporate performance: A stakehold perspect. Sustainability 2015, 7, 4997-5026.
[2] *Carter, C.R., Rogers, D.S.* A framework of sustainable supply chain management: Moving toward new theory. Int. J. Phys. Distrib. Logist. Manag. 2008, 38, 360-387.
[3] *Hervani, A.A., Helms, M.M., Sarkis, J.* Performance measurement for green supply chain management. Benchmarking 2005, 12, 330-353.
[4] *Bhattacharya, A., Mohapatra, P., Kumar, V., Dey, P.K., Brady, M., Tiwari, M.K., Nudurupati, S.S.* Green supply chain performance measurement using fuzzy ANP-based balanced scorecard: A collaborative decision-making approach. Prod. Plan. Control 2014, 25, 698-714.
[5] *Shen, L., Olfat, L., Govindan, K., Khodaverdi, R., Diabat, A.* A fuzzy multi criteria approach for evaluating green supplier's performance in green supply chain with linguistic preferences. Resour. Conserv. Recycl. 2013, 74, 170-179.
[6] *Uygun, Ö., Dede, A.* Performance evaluation of green supply chain management using integrated fuzzy multi-criteria decision-making techniques. Comput. Ind. Eng. 2016, in press.
[7] *Richardson, P.* Fitness for the Future: Applying Biomimetics to Business Strategy. Ph.D. Thesis, Bath University, Bath, UK, 2010.
[8] *Li, Y.* Application of GA and SVM to the performance evaluation of supply chain. Comput. Eng. Appl. 2010, 46, 246-248.
[9] *Fan, X., Zhang, S., Wang, L., Yang, Y., Hapeshi, K.* An evaluation model of supply chain performances using 5DBSC and LMBP neural network algorithm. J. Bionic Eng. 2013, 10, 383-395.
[10] *Xi, Y.F., Wang, C., Nie, X.X.* Study of the evaluation methods of supply chain performances using fuzzy neural network. Inf. J. 2007, 9, 77-79
[11] *Seeley, T.D.* The Wisdom of the Hive: The Social Physiology of Honey Bee Colonies, Harvard University Press: Cambridge, UK, 1996.
[12] *Martens, D., De Backer, M., Haesen, R., Vanthienen, J., Snoeck, M., Baesens, B.* Classification with ant colony optimization. IEEE Trans. Evol. Comput. 2007, 11, 651–665.
[13] Seeley, T.D. The Wisdom of the Hive: The Social Physiology of Honey Bee Colonies, Harvard University Press: Cambridge, UK, 1996.
[14] *Gabbert, P., Brown, D.E., Huntley, C.L., Markowicz, B.P., Sappington, D.E.* A system for learning routes and schedules with genetic algorithms. In Proceedings of the Fourth International Conference on Genetic Algorithms (ICGA'91), Santt Diego, CA, USA, 13–16 July 1991, pp. 430-436.
[15] *Allaoui, H., Guo, Y., Choudhary, A., Bloemhof, J.* Sustain able agro-food supply chain design using two-stage hybrid multi-objective decision-making approach. Computers & Operations Research, 2018, 89(1): 369-384.
[16] *Guo, Y., Hu, F., Allaoui, H., Boulaksil, Y.* A distributed approximation approach for solving the sustainable supply chain network design problem. International Journal of Production Research, 2019, 57(11): 3695-3718.

[17] *Muhammad Amad saeed, Wolfgang, Kersten.* Supply chain sustainability performance indicators- a content analysis based on published standards and guidelines. Logistic Research, 2017, 10-12.

[18] *Benjaafar, S., Li, Y., & Daskin*, M. Carbon footprint and the management of supply chains: insights from simple models. IEEE Transactions on Automation Science & Engineering, 2012, 10(1), 99-116.

[19] *Xie, G.*. Modeling decision processes of a green supply chain with regulation on energy saving level. Computers & Operations Research, 2015, 54, 266-273.

[20] *Shao, K.*; *Yang, J.* Study on the performance evaluation of green supply chain Based on the balance scorecard and fuzzy theory. In Proceedings of the 2nd IEEE International Conference on Information Management and Engineering (ICIME), Chengdu, China, 16–18 April 2010; pp. 242-246.

[21] *Coskun, S., Ozgur, L., Polat, O., & Gungor,* A. A model proposal for green supply chain network design based on consumer segmentation. Journal of Cleaner Production, 2016, 110, 149-157.

[22] *Rezaee, A., Dehghanian, F., Fahimnia, B., & Beamon, B*. Green supply chain network design with stochastic demand and carbon price. Annals of Operations Research, 2015, 250(2), 1-23.

[23] *Zhalechian M, Tavakkoli-Moghaddam R, Zahiri B, et al*. Sustainable design of a closed-loop location-routing-inventory supply chain network under mixed uncertainty. Transportation Research Part E Logistics & Transportation Review, 2016, 89:182-214.

[24] *Roy, A., Sana, S. S., Chaudhuri, K.* Optimal Pricing of Competing Retailers Under Uncertain Demand - A Two Layer Supply Chain Model. Annals of Operations Research, 2018, 260 (1-2): 481-500.

[25] *K. Shaw, R. Shankara, S.S Yadava, L.S. Thakurb.* Supplier selection using fuzzy AHP and fuzzy multi-objective linear programming for developing low carbon supply chain Expert Syst Appl, 2012, 39-9, 8182-8192

[26] *Hosseini, S., & Khaled, A. A.* A hybrid ensemble and AHP approach for resilient supplier selection. Journal of Intelligent Manufacturing, 2017, 1-22.

[27] *Aksen, D., Piyade, N. and Aras, N.* The budget constrained r-interdiction median problem with capacity expansion. Central European Journal of Opearations Research, 2009, 18(3), 269-291.

[28] *Vencheh A. H., Matin R K, Kajani M T*. Undesirable factors in efficiency measurement. Applied Mathematics & Computation, 2005, 163(2): 547-552.

[29] *Liou, J. J. H., J., Zavadskas, E. K., & Tzeng, G. H*. New hybrid copras-g madm model for improving and selecting suppliers in green supply chain management. International Journal of Production Research, 2016, 54(1), 114-134.

[30] *Roy, A., Sana, S. S., & Chaudhuri, K.* Optimal pricing of competing retailers under uncertain demand-a two layer supply chain model. Annals of Operations Research, 2015, 1-20.

[31] *Wang, G., Lei, L., & Lee, K*. Supply chain scheduling with receiving deadlines and non-linear penalty. Journal of the Operational Research Society, 2015, 66(3), 380-391.

[32] *Oliveira, C., Coelho, D., & Antunes, C. H*. Coupling input–output analysis with multiobjective linear programming models for the study of economy–energy–environment–social (e3s) trade-offs: a review. Annals of Operations Research, 2014, 1-32.

[33] *Sarkis, J.* Policy insights from a green supply chain optimisation model. International Journal of Production Research, 2015, 53(21), 6522-6533.

[34] *Kannan, D.* Evaluation of green manufacturing practices using a hybrid model combining danp with promethee. International Journal of Production Research, 2015, 53(21), 6344-6371.

[35] *Xie, G*. Modeling decision processes of a green supply chain with regulation on energy saving level. Computers & Operations Research, 2015, 54, 266-273.
[36] *Mirhedayatian, S. M., Azadi, M., & Saen, R. F.* A novel network data envelopment analysis model for evaluating green supply chain management. International Journal of Production Economics, 2014, 147(1), 544-554.
[37] *Barbosa-Póvoa, A. P., Silva, C. D., Carvalho, A.* Opportunities and Challenges in Sustainable Supply Chain: An Operations Research Perspective. European Journal of Operational Research, 2018, 268: 399–431.
[38] *Sarkis, J.* Policy insights from a green supply chain optimisation model". International Journal of Production Research, 2015, 53(21), 6522-6533.
[39] *Boonsothonsatit, K., Kara, S., Ibbotson, S., & Kayis, B*. Development of a generic decision support system based on multi-objective optimisation for green supply chain network design (goog). Journal of Manufacturing Technology Management, 2015, 26(7).
[40] *Slimani, I., Farissi, I. E., & Achchab, S*. "Coordination by sharing demand forecasts in a supply chain using game theoretic approach". IEEE International Colloquium on Information Science and Technology, 2016, 122-127.
[41] *Validi S, Bhattacharya A, Byrne P J.* A case analysis of a sustainable food supply chain distribution system—A multi-objective approach. International Journal of Production Economics, 2014, 152(4):71-87.
[42] *Roghanian, E., & Pazhoheshfar, P*. An optimization model for reverse logistics network under stochastic environment by using genetic algorithm. Journal of Manufacturing Systems, 2014, 33(3), 348-356.
[43] *Jiang, D., Li, H., Yang, T., & Li, D*. Genetic algorithm for inventory positioning problem with general acyclic supply chain networks. European J of Industrial Engineering, 2016, 10(3), 367.
[44] *Jamshidi, R., Ghomi, S. M. T. F., & Karimi, B*. Multi-objective green supply chain optimization with a new hybrid memetic algorithm using the taguchi method". Scientia Iranica, 2012, 19(6), 1876-1886.
[45] *Carlucci, S., Cattarin, G., Causone, F., & Pagliano, L.* Multi-objective optimization of a nearly zero-energy building based on thermal and visual discomfort minimization using a non-dominated sorting genetic algorithm (nsga-ii). Energy & Buildings, 2015, 104, 378-394.
[46] *Dotoli, M., Epicoco, N., Falagario, M*. A Fuzzy Technique for Supply Chain Network Design with Quantity Discounts. International Journal of Production Research, 2017, 55 (7): 1862-1884.