A Coevolutionary Estimation of Distribution Algorithm for Group Insurance Portfolio

Wen Shi, *Student Member, IEEE*, Wei-Neng Chen, *Senior Member, IEEE*, Sam Kwong, *Fellow, IEEE*,
Jie Zhang, Hua Wang, *Senior Member, IEEE*, Tianlong Gu, Huaqiang Yuan, and Jun Zhang, *Fellow, IEEE*

Abstract—With the rapid development of the insurance industry, more diverse insurance products are produced for consumers. Insurance portfolio problems have received increasing attention. While most studies focus on insurance portfolio problem for a single insured, insurance portfolio problems for a specific group of insured are even more intricate but little attention has been paid to. In this article, we propose a group insurance portfolio model for investment allocation of several insurance policies so that the total payout of the whole group can be maximized. The statistical average value of each parameter is considered in the model to approximate the expectation payout of the group insurance portfolio problem. To solve this problem, a coevolutionary estimation of distribution algorithm (EDA) utilizing the divide-and-conquer strategy is proposed. First, as the payout of each insured under a certain portfolio plan can be calculated separately, the proposed approach decomposes the group insurance portfolio problem into several single-insured insurance portfolio problems. In this way, the dimension of the optimization problem becomes lower compared to the original problem. An adaptive EDA is proposed to optimize the portfolio plan of each insured independently. Second, the group insurance portfolio problem remains a nonseparable problem since the investment amount of each insured is limited by the total investable amount of the whole group. A particle swarm optimization algorithm is adopted to cooperate with the EDA to optimize the proportion of allocation to each insured. The proposed algorithm is verified on various scenarios. The experimental results validate that the proposed approach is effective for the group insurance portfolio problem.

Index Terms—Cooperative coevolution (CC), estimation of distribution algorithm (EDA), insurance portfolio, particle swarm optimization (PSO).

I. INTRODUCTION

NOWADAYS, more and more people consider insurance as an important part of their investment plan for risk aversion. Making a proper selection of insurance policies becomes an important issue. However, due to the rapid development of the insurance industry, a great number of insurance policies have emerged. Facing these massive policies, it is quite difficult to make a proper choice.

Fruitful of literature about insurance investment have been published. Some of them studied the insurance investment of a single insured while the others studied the insurance investment of a whole group. Early studies [1]–[5] considered only life insurance with death compensation for a single insured. In these researches, the payout of life insurance is considered proportional to the premium. Such simplified models may have some limitations in practical applications. On the one hand, the traditional models only consider a single life insurance. While there are an enormous number of insurance policies to be chosen in real-life insurance market. Therefore, these models cannot fulfill the market requirement due to the scarcity of the kinds of policies. On the other hand, traditional models regard the payout in proportion to the premium. While the relationship between the premium and payout is much more complicated in real-life insurance policies. Therefore, these models are not well consistent with real-life policies.

To be more practical, some literature proposed the insurance portfolio problem by studying the insurance investment strategy among several insurance products. For example, Mousa *et al.* [6] proposed an insurance portfolio model composed of several life insurance products. In our previous study [7], a data-driven single-insured insurance portfolio model is proposed, which applied massive data to express the relationship between the payout and premium. However, the above models only studied the insurance portfolio problem for a single insured.

In some situations, policyholders need to consider insurance investment plan for a group of people [8]–[10]. For example, a mother will consider insurance portfolio plan for

the whole family instead of only herself. An employer will consider insurance portfolio plan for all employees in the company instead of only himself. In these cases, the objectives of the policyholders are to optimize the total payout of the whole group instead of a single person. However, a group always consists of insureds with different ages, jobs, and health conditions, which make it necessary to consider different insurance policies for different insureds. Meanwhile, the complex relationships among insureds in the group will influence the investment tendencies. Thus, in order to alleviate the risk of the group and support the elderly or invalid, it is important to study the group insurance portfolio problem. Huang *et al.* [8] took the whole family into consideration and applied different wealth accumulation functions after the death or retirement of parents. Their work was then followed by Kwak *et al.* [9], which considered the utility function of parents and children separately. After that, Bruhn and Steffensen [10] considered the cases of *n* insured and treated death as a random process. In these articles, only one type of life insurance is considered with a payout proportional to the premium.

This article proposes a data-driven group insurance portfolio model, which is a better approximation of reality than traditional models. A novel coevolutionary method is devised to approximate the optimal solution.

To solve group insurance portfolio problem consisting of multiple insurance policies, new challenges are posed from two aspects. The first one is the establishment of the optimization model. In the group insurance portfolio problem, several types of endowment policies and hospitalization policies are taken into consideration. Traditional methods cannot fully model the relationship between the payout and the premium. Actually, since many factors (e.g., investable amount, age structure, investment duration, etc.) should be taken into account during the calculation, it is quite hard to deduce a mathematical model for the payout and the premium. The second one is how to devise the optimization algorithm. The insurance portfolio plan for a group consists of the asset allocation to each insured in the group, the investment amount of each insured to each endowment policy, and the choice of hospitalization policies of each insured. These variables are in different solution space. The diversity of the solution space makes the problem so complex that its mathematical properties cannot fulfill the premises of using analytic methods. It is difficult to apply traditional optimization methods to address the model [7].

Regarding the issue above, we study the group insurance portfolio problem in the following two aspects.

1) First, a data-driven model is established to describe the premium and the payout using the accumulated historical data. Inspired by the data-driven model for a single insured proposed in our previous work [7], a modified model for the group problem is proposed. In the group model, the premium and payout for each insured is different, which is based on their individual situation. The statistical average value of each parameter is considered in the model to approximate the expectation payout of the problem. The objective is to maximize the total payout of the whole group instead of a single insured.

2) Second, since the model is so complicated that the optimal solution is difficult to gain by mathematical calculation, an effective evolutionary algorithm, estimation of distribution algorithm (EDA) [11], is applied to deal with this model. Compared to the case of a single insured, the group case becomes more intricate. With the increase of insured considered in the portfolio problem, the possible combination of the investment on insurance policies increases exponentially. The optimization problem changes from a low-dimensional problem to a high-dimensional problem. In order to improve the performance of EDA in large-scale optimization, a novel coevolutionary method is proposed in this article to make the optimization algorithm more effective.

Inspired by the cooperative coevolution (CC) method [12], this article proposes a coevolutionary EDA (CEDA) to adapt to the characteristics of the group insurance portfolio problem. On the one hand, since the total premium is limited, the portfolio plan of each insured will affect others. The impact of each insured's portfolio plan on overall revenue is reflected in the association of total limits. In terms of this constraint, the problem is indivisible. On the other hand, for each insured in the group, the insurance portfolio problem can be seen as a single-insured insurance portfolio problem. In such a problem, the payout of the portfolio plan is evaluative. Therefore, in terms of the objective function, the problem is independent to some extent.

In order to adapt to the above two characteristics, a special coevolutionary method is proposed. To ensure the indivisibility, a coevolutionary process is applied to optimize the allocation of the investable asset of each person with the portfolio plan simultaneously by a classic particle swarm optimization (PSO) algorithm. Due to the independence of the problem, the cooperative process can be left out in this problem. An adaptive EDA (AEDA), which has been validated to be effective for single-insured insurance portfolio problem [7], is applied to optimize each subcomponent independently. By this mean, the optimization process can be carried out effectively.

The remainder of this article is organized as follows. Section II introduces the backgrounds of insurance investment model, CC method, and EDA. In Section III, a data-driven model is proposed to approximate the expectation payout of the group insurance portfolio problem. Section IV elaborates on the proposed CEDA in detail. Section V presents the experimental results to validate the effectiveness of the proposed CEDA. Finally, the conclusions and summaries of the article are provided in Section VI.

II. BACKGROUND

A. Insurance Investment Model

1) Single-Insured Model: Numerous scholars have investigated the insurance investment problem in recent years. Yaari [1] first took the uncertain life time of the consumer into consideration and studied the purchase of life insurance. After that, Hakansson [2] followed this idea and considered the investment of insurance as a part of the wage earner in discrete cases. Richard [3] extended this idea to the continuous cases. Then, Pliska and Ye [4] proposed the investment problem with

a risk-free asset and a life insurance policy. Ye [5] followed this study and added a risky asset into the investment problem. Duarte *et al.* [13], [14] further changed one risky asset to several risky assets to carry on the research. Stabile [15] focused on the investment of annuity and made a research to find out the optimal timing of annuity purchase for the wage earner. Pirvu and Zhang [16] modified the research with stochastic income function and the risk function of the stocks. Kronborg and Steffensen [17] brought about the consideration for surrender option. In this model, the wage earner will sell the life insurance when his wealth is lower than a certain extent. Shen and Wei [18] considered the parameters in the model as functions related to time t. In most of the researches, the investment of insurance is viewed as a part of the whole investment plan. The wealth accumulation function at t of the policyholder is described as [3]–[5]

$$dM(t) = (W(t) - c(t) - I(t) + \beta(t)M(t))dt$$
$$+ \sum_{i=1}^{n} dS_i(t), t < \tau \qquad (1)$$
$$M(t) = M(t^-) + Z(t), t = \tau. \qquad (2)$$

Here, $M(t)$ denotes the vendibility of the family's assets; $W(t)$ is the revenue rate of the wage earner in the family; $c(t)$ is the instantaneity consumption rate of the family; $I(t)$ is the premium of life insurance; $\beta(t)$ is the constant instantaneous expected percentage change in price of the riskless security; $\tau = \min(\tau_0, T)$ is the earlier time between the time of retirement T and time of death τ_0; and $Z(t)$ is the death compensation of the life insurance, which is also the payout of life insurance in this model. $S_i(t)$ is the vendibility of each risky asset, which is generated by a nonstationary geometric Brownian motion process

$$dS_i(t) = w_i(t)M(t)\left[\alpha_i(t)dt + \sigma_i(t)dq_i(t)\right] \qquad (3)$$

where $w_i(t)$ is the fraction of vendibility invested in risky asset i [i.e., $w_i(t)M(t) = S_i(t)$]; $q_i(t)$ is a standard normal random variable with mean zero and variance one; $\alpha_i(t)$ is the constant instantaneous expected percentage change in price; and $\sigma_i^2(t)$ is the constant instantaneous variance.

Death compensation $Z(t)$ in (2) only pays out when the policyholder dies or retires. $Z(t)$ is considered proportional to the premium, i.e.,

$$Z(t) = \frac{I(t)}{\eta(t)} \qquad (4)$$

where $\eta(t)$ denotes the premium-payout ratio.

Besides, a risk-averse utility function is utilized in the models to measure the satisfaction gained from consumption. This function only considers the impact of personal mental state on the degree of risk aversion, regardless of the influence of the possession of wealth to risk aversion, i.e.,

$$u(c) = \frac{1}{1 - \gamma} c^{1-\gamma} \qquad (5)$$

where $\gamma > 0$ is the coefficient of relative risk aversion.

However, since more and more insurance products have arisen, only considering a single insurance policy is not realistic enough. Therefore, an insurance investment model [6],

which has taken several types of life insurance policies into consideration, has been proposed. In this model, the wealth accumulation function is the same as the previous models, but the death compensation $Z(t)$ is the sum of policies [6]

$$Z(t) = Z_1(t) + \cdots + Z_k(t) = \frac{I_1(t)}{\eta_1(t)} + \cdots + \frac{I_k(t)}{\eta_k(t)} \qquad (6)$$

where $Z_i(t)$ is the death compensation of the ith life insurance policy; $I_i(t)$ is the premium of the ith life insurance policy; and $\eta_i(t)$ is the premium-payout ratio of the ith life insurance policy.

In the above model, different policies are with different premium-payout ratios at different time. However, it is obvious that the best investment plan is to choose the policy with the highest premium-payout ratios. Even though different policies are considered, this model oversimplifies the payout mechanism and thus, cannot suit the practical market well.

Actually, few companies simply offer life insurance policies with only death compensation in reality. Endowment policies [19] with the function of savings are much more popular. Traditional models cannot fully express the relationship between the payout and the premium.

In our previous work [7], a data-driven model is proposed to solve the insurance investment problem considered several types of hospitalization policies in addition to endowment policies. The statistical value of the total payout at a specific time t is approximated as

$$J(t; X, \beta) = \left[1 - p(t)\right]V(t) + p(t)Z(t) + C(t) \qquad (7)$$

where X denotes the investment plan of endowment policies and β is the selected hospitalization policy. $V(t)$ is the cash value, which cannot be extracted by the policyholder until the termination of the policy; $Z(t)$ is the death compensation, which can be gain after the death of the insured; $C(t)$ is the cash the policyholder can be disposed of at will; and $p(t)$ is the mortality rate of the insured.

In this data-driven model, the historical data of the death toll are utilized to estimate the mortality rate of the insured. The historical incidence rate is utilized to estimate the medical cost and calculate the optimal solution of β. Historical data of the payout are utilized to evaluate the cash value and death compensation of each endowment policy. The historical data of the income are utilized to evaluate the investable amount.

By using these huge amounts of historical data, an approximation of the expectation payout of the insurance portfolio problem can be gained and an individualized insurance portfolio plan can be optimized precisely for a specific insured.

2) Group Model: However, in some situations, the policyholders care about how to get the most payout for a whole group instead of a single person when making an insurance portfolio plan.

Many researchers discover this phenomenon and study the insurance investment model for the whole group [8]–[10]. Huang *et al.* [8] considered the investment problem for the whole family instead of a single policyholder. Different wealth accumulation functions are considered before and after the death or retirement of parents.

Before the death or retirement of parents, i.e., $t < \tau$, since there are still annual incomes from parents, the wealth accumulation function is the same as (1). After the death or retirement of parents, i.e., $t > \tau$, since there are no annual incomes from parents, the wealth accumulation function changes to

$$dM(t) = -c(t)\mathrm{dt} + d\alpha(t) + M(t)\beta(t)\mathrm{dt}. \tag{8}$$

A short-term life insurance is bought for the parents. At the time when parents die or retire, i.e., $t = \tau$, the payout of the life insurance is calculated as (4). Here, both parents and children use the risk-averse utility function given in (5).

Based on this work, Kwak *et al.* [9] further considered different utility functions for children and parents

$$u_c(c) = \frac{(c - R_c)^{1-\gamma_c}}{1 - \gamma_c} \tag{9}$$

$$u_p(c) = \frac{(c - R_p)^{1-\gamma_p}}{1 - \gamma_p} \tag{10}$$

where γ_c is the coefficient of relative risk aversion for children and γ_p is the coefficient of relative risk aversion for parents.

Although these studies just consider the investment in insurance for a single insured, the idea that takes a whole group into consideration and maximizes the payout of the whole group is very thought provoking.

Inspired by these studies, we propose a data-driven group insurance portfolio model, which considers the insurance portfolio problem for a group of insured to approximate the expectation payout of the problem. When the number of considered insured increases, more insurance policies can be chosen and the dimensions of the problem rise correspondingly, which brings a new challenge to solve the problem.

B. Cooperative Coevolution Method

CC methods [12], [20], [21] are widely used to improve the performance of evolutionary algorithms in large-scale optimization problems [22], [23].

The core idea of the CC method is to decompose a high-dimensional problem into several lower subcomponents so that it is much easier to deal with each subcomponent separately. The pseudocode of the CC method is given in Algorithm 1 and the detailed procedure consists of the following three parts.

1) Decomposition: As shown in Algorithm 1, the CC method first initializes a population consisted of a group of candidate solutions. Then, these solutions are decomposed into several subcomponents with lower dimension. Early studies applied simple strategies without the consideration of problem characteristics, such as one-dimensional-based strategy [24], splitting in half strategy [25], and random grouping strategy [26]. Then, problem-based-decomposition strategies were proposed, which can be classified into two categories: 1) fixed decomposition strategies and 2) dynamic decomposition strategies. The subcomponents obtained by fixed decomposition strategies will not change once determined throughout the evolution process while the subcomponents obtained by dynamic decomposition strategies keep changing during the iterations. Differential grouping (DG) algorithm [27] is a popular fixed decomposition strategy. It decomposed the problems based on

---

Algorithm 1 Cooperative Coevolution Method

1: **Initialization**
2: Decompose the variables into m sub-components
3: **while** termination condition not satisfied **do**
4:     **for** $i = 1$ to m **do**
5:         Optimize the i-th sub-component
6:     **end for**
6: **end while**

Algorithm 2 Estimation of Distribution Algorithm

1: **Initialization**
2: **while** termination condition not satisfied **do**
3:     Choose the elite solutions
4:     Establish probability model with these solutions
5:     Generate new population with the probability model
6: **end while**

---

interaction between decision variables. To make up for the limitation of DG that only direct interaction between variables can be identified, an extended DG (XDG) algorithm [28] was proposed. XDG algorithm can identify not only direct interaction between variables but also indirect interaction. To increase the accuracy of DG, a global DG (GDG) algorithm [29] was proposed. Global information is maintained in GDG algorithm so that the accuracy can be increased. Multilevel CC (MLCC) [30] method is a popular dynamic decomposition strategy. MLCC adaptively adjusted the size of the subcomponent. Delta grouping [31] is another popular dynamic decomposition strategy. It identified interaction variables by the change amount in each decision variable in each generation and decomposed the problem based on it.

2) Optimization: After that, each subcomponent evolves independently using a certain evolutionary algorithm. Potter and Jong [12] first applied genetic algorithms (GAs) [32] as the optimizer of the subcomponents and found it effective. Then, researchers [26], [33]–[35] applied other evolutionary algorithms, such as the differential evolution (DE) algorithm [36] and PSO [37] as the optimizer of the subcomponents.

3) Cooperation: Since the fitness value cannot be calculated using only one subcomponent, cooperation occurs during the fitness evaluation process. In the optimization process, the solution with global best fitness value is archived. In the evolution process of each subcomponent, except for the dimensions to be optimized, other dimensions keep the same as the global best solution. That is, denoted the global best solution as $X_{\text{best}} = (X_1^{\text{best}}, \ldots, X_m^{\text{best}})$, to optimize the ith subcomponent in the gth generation $X_{i,g}$, solution $X = (X_1^{\text{best}}, \ldots, X_{i-1}^{\text{best}}, X_{i,g}, X_{i+1}^{\text{best}}, \ldots, X_m^{\text{best}})$ is utilized to evaluate the fitness value.

C. Estimation of Distribution Algorithm

Due to its great global search ability, EDA [38] is widely used in various fields [39]–[41]. As shown in Algorithm 2, it first randomly generates the initial population. Then, the promising solutions with relatively better fitness values are

TABLE I
NOTATION

$t_0{}^k$	The initial age of the k-th insured in the group
I_1	The initial investable amount of the whole group
$I_2{}^k(t)$	The increment of the investable amount of the k-th insured at age t
n_1	The number of endowment policies
n_2	The number of hospitalization policies
n	The number of insured in the group
m_i	The number of payment period for the i-th endowment policy
$x_{ij}{}^k$	The premium of the j-th payment period of the i-th endowment policy for the k-th insured in the group
β^k	The kind of hospitalization policy the k-th insured invest in ($\beta^k = 0$ implies does not buy any hospitalization policy)
q_{ij}	The latest time of purchase for the j-th payment period of the i-th endowment policy
l_{ijt}	The minimum amount for purchase of the j-th payment period of the i-th endowment policy at age t
T_1	The guarantee period of the endowment policies
T_2	The latest time to buy hospitalization policies
T_3	The duration to be considered
$\eta_{1ij}{}^k(t_1, t)$	The increase rate of the cash value of the i-th endowment policy with the j-th payment period for the k-th insured at age t with purchase at age t_1
$\eta_{2ij}{}^k(t_1, t)$	The death compensation rate of the i-th endowment policy with the j-th payment period for the k-th insured at age t with purchase at age t_1
$\delta_\beta{}^k(t)$	The premium of hospitalization policy β for the k-th insured at age t
$p_s{}^k(t)$	The incidence rate of the k-th insured for diseases of degree s at age t
$p^k(t)$	The mortality rate of the k-th insured at age t
z_s	The medical expense of diseases at degree s
$V^k(t)$	The cash value of the k-th insured at time t
$Z^k(t)$	The death compensation of the k-th insured at time t
$C(t)$	The total disposable cash for the whole group at time t
$V_{ij}{}^k(t)$	The cash value of the i-th endowment policy with the j-th payment period for the k-th insured at time t
D_k	The total investment amount of the k-th insured
D	The total investment amount of the whole group.

selected. A probability model is established with these solutions by the statistical learning method to describe the distribution of the candidate solutions in the search space. Then, a new population is randomly generated according to this probability model. The above processes are executed repeatedly to achieve the evolution of the population.

III. GROUP INSURANCE PORTFOLIO MODEL

This study considers the insurance portfolio problem for a group consists of several persons. The policyholder is the person who makes the decision on the portfolio plan and the insureds are all of the persons in the group. The notation used in the model is provided in Table I for better understanding.

A. Problem Description

In the case considered in this study, an insurance portfolio plan should be made for a group consisting of n persons to maximize the total payout for duration T_3. There are n_1 kinds of endowment policies and n_2 kinds of hospitalization policies for each insured to choose, where the ith endowment policy has m_i kinds of payment period. The latest time to buy the ith endowment policy with the jth payment period is q_{ij} and the investment amount for an insured of age t is at least l_{ijt}. The premium and combinations of endowment policies

can be decided by the policyholder discretionarily as long as the minimum requirement is satisfied. Besides, each person in the group should at least make an investment of a certain amount to fulfill the basic demand. The guarantee period of the endowment policies is T_1. After T_1, if the policyholder is still alive, he can get the same amount of payout as the death compensation at time T_1.

The latest time to buy hospitalization policies is T_2. Each insured can only choose one hospitalization policy to invest. Once he decides which kind of hospitalization policy to invest, he cannot change to another one and the premium will not change. The premium of hospitalization policy keeps the same amount as the investment begins. To pay the premium of hospitalization policy, the cash value in endowment policies will be deducted accordingly by insurance company to cover it.

According to serious degrees, diseases in concern are categorized into n_2 groups and indexed in descending order. So that the diseases in the first group are the most serious ones while the diseases in the n_2th group are the least serious ones. The coverage scope of hospitalization policy increases as the premium increases. The coverage scope of the first policy is the smallest so that the premium is the lowest. It can just cover the diseases in the first group. As for the second policy, the coverage scope is the diseases in the first two groups. The coverage scope of the rest policies can be deduced from this. For the n_2th policy, which is the most expensive one, all of the diseases in concern are covered. The probability of having the diseases in degree s at age t for the kth insured is denoted as $p_s{}^k(t)$ and the mortality rate is denoted as $p^k(t)$, which is different for each insured depending on his health condition.

Given the detailed information of each insured in the group, this study aims to optimize the insurance portfolio plan for the whole group to maximize the statistical approximation value of the expectation total payout.

B. Model Construction

To consider the insurance portfolio plan of a whole group consisting of n persons, the target is to maximize the total payout while the total disposable income is fixed.

Extended from the single-insured insurance portfolio model proposed in [7], the approximation model of the expectation payout of group insurance portfolio problem is formulated as

$$J\left(X^k, \beta^k\right) = \sum_{t=1}^{T_3} \sum_{k=1}^{n} \left(1 - p^k\left(t + t_0^k - 1\right)\right)$$
$$\times \left[\left(1 - p^k\left(t + t_0^k\right)\right)V^k(t) + p^k\left(t + t_0^k\right)Z^k(t)\right]$$
$$+ C(t) \qquad (11)$$

where X^k is an irregular matrix where element $x_{ij}{}^k$ is the investment amount for each endowment policy for the kth insured. T_3 is the duration to be taken into account. β^k is the kind of hospitalization policy the kth insured invests in. $p^k(t)$ denotes the mortality rate of the kth insured. $V^k(t)$ is the total cash value of the kth insured and $Z^k(t)$ is the death compensation of the kth insured. $C(t)$ is the total disposable cash for the whole group.

$V^k(t)$ in (11) is calculated by summing up all of the cash value of each endowment policy for the kth member, i.e.,

$$V^k(t) = \sum_{i=1}^{n_1} \sum_{j=1}^{m_i} V_{ij}^k(t) \tag{12}$$

where n_1 is the number of endowment policies and m_i is the number of payment period for the ith endowment policy. $V_{ij}^k(t)$, which is determined by (13), is the cash value of each endowment policy for the kth insured at time t

$$V_{ij}^k(t) = \begin{cases} x_{ij}^k \eta_{1ij}^k(t_0^k, t_0^k + t), & \text{if } t = 1 \\ V_{ij}^k(t-1) \\ \quad + x_{ij}^k \left[\eta_{1ij}^k(t_0^k, t_0^k + t) \right. \\ \quad \left. - \eta_{1ij}^k(t_0^k, t_0^k + t - 1) \right], & \text{if } 1 < t < \xi^k \\ 0, & \text{if } t \geq \xi^k \end{cases} \tag{13}$$

Here, $\xi^k = \min\{T_1 - t_0^k, T_3\}$ is the latest time that the kth insured can get the cash value. $\eta_{1\,ij}^k(t_1, t)$ is the increase rate of cash value for each endowment policy at age t with purchase at age t_1 for the kth insured. These cash values are not disposable except for the payment of the premium of hospitalization policy. Note that for each insured, only the cash value of himself can be used to pay his hospitalization premium. Cash values are taken out in descending order to pay the premium until it is paid off. If the premium cannot be covered by the cash values [i.e., $\delta^k(t) > V^k(t)$] for any insured k at any time t, the hospitalization policy is considered unaffordable. In this case, this portfolio plan is defined as infeasible and will not be considered any more. For endowment policy whose cash values are taken out to pay the premium, its cash values $V_{ij}^k(t)$ are reduced correspondingly.

Similarly, denoted the death compensation rate of each endowment policy for the kth insured at age t with purchase at age t_1 as $\eta_{2\,ij}^k(t_1, t)$, the death compensation $Z^k(t)$ is calculated by summing up all of the death compensations of endowment policies for the kth insured, i.e.,

$$Z^k(t) = \sum_{i=1}^{n_1} \sum_{j=1}^{m_i} x_{ij} \eta_{2ij}^k(t_0^k, t_0^k + t). \tag{14}$$

As for the disposable cash in hand $C(t)$ for the whole group, it is calculated by the sum of the balance at time $(t-1)$ and the disposable income for insurance investment of the whole group $I(t)$, and subtracts the medical expenses that are not covered by the selected hospitalization policy of each person. Denoted the medical expense for the diseases in the sth group as z_s, $C(t)$ is calculated as

$$C(t) = \begin{cases} I_1, & \text{if } t = 0 \\ C(t-1) + I(t) - \sum_{k=1}^{n} \\ \quad \times \left(\sum_{i=1}^{n_1} \sum_{j=1}^{m_i} v_{ij}^k(t) x_{ij}^k \right. \\ \quad \left. - \sum_{s=1}^{n_2 - \beta^k} p_s^k(t + t_0^k) z_s^k(t) \right), & \text{otherwise} \end{cases} \tag{15}$$

$$I(t) = \sum_{k=1}^{n} I_2^k(t + t_0^k) \tag{16}$$

where $v_{ij}^k(t) \in \{0, 1\}$ is an indicator of whether or not the kth member invests in each endowment policy. $p_s^k(t)$ is the probability that the kth insured will suffer from the diseases in the sth group. $I_2^k(t)$ is the income of the kth insured at age t.

Besides, each insured in the group should invest in a minimum amount of endowment policy to guarantee the basic need. Therefore, a penalty mechanism given in (17) is implemented to support this demand

$$J'\left(X^k, \beta^k\right) = J\left(X^k, \beta^k\right) \prod_{k=1}^{n} v_k \tag{17}$$

$$v_k = \begin{cases} 0.9 + \frac{D_k}{D}, & \text{if } \frac{D_k}{D} < 0.1 \\ 1, & \text{if } \frac{D_k}{D} \geq 0.1. \end{cases} \tag{18}$$

Here, the minimum amount for each insured to invest is set to be 10% of the whole investment amount. v_k is the penalty rate. D_k is the total investment amount of the kth insured while D is the total investment amount of the whole group. $J'(X^k, \beta^k)$ is the final objective function, which we need to maximize.

C. Parameters Estimation

In the proposed data-driven model, a great many of parameters are needed to ensure that the optimization process works properly.

For insurance company, the following data should be given as parameters in the model.
1) $\eta_{1ij}^k(t_0^k, t_0^k + t)$: The increase rate of cash value for each insured for each endowment policy.
2) $\eta_{2ij}^k(t_0^k, t_0^k + t)$: The death compensation rate for each insured for each endowment policy.
3) m_i: The number of payment period for each endowment policy.
4) q_{ij}: The latest time to purchase each endowment policy.
5) l_{ijt}: The minimum investment amount required by the insurer for each endowment policy.
6) T_1: The guarantee period of the endowment policies.
7) $\delta_\beta^k(t)$: The premium of hospitalization policy when the investment is made at age t.
8) T_2: The age limitation of hospitalization policies.

For the insured, the following data should be decided by the policyholder as parameters in the model.
1) I_1: The initial investment amount of the whole group.
2) $I_2^k(t)$: The increment of the investable amount for each insured at time t.
3) t_0^k: The beginning age to purchase insurance for each insured.
4) T_3: The duration to be considered.
5) n_1: The number of available endowment policies.
6) n_2: The number of available hospitalization policies.
7) n: The number of insured considered in the group.

Except for the above parameters, there are other parameters that cannot be obtained directly. To gain these parameters, statistical analysis should be done.
1) $p^k(t)$: The mortality rate of each insured

$$p^k(x + \tau) = \frac{s^k(x + \tau + 1) - s^k(x + \tau)}{s^k(x)}. \tag{19}$$

Here, $s^k(.)$ denotes the number of survivors which can be estimated by the age distribution over the people with similar conditions to each insured. To gain these age distributions, historical data should be collected to do analysis.

2) $p_s^k(t)$: The incident rate of each insured for diseases of each degree

$$p_s^k(x + \tau) = \frac{i_s^k(x + \tau)}{s^k(x + \tau)}. \tag{20}$$

Here, $i_s^k(.)$ denotes the number of patients suffer from diseases of degree s. To conduct the above derivation, $i_s^k(.)$ and $s^k(.)$ should be gained by the estimation of the incident age distribution of patients with similar conditions to the kth insured. Historical data are also necessary to achieve such estimation.

3) z_s^k: The medical expense of diseases at each degree for each insured

$$z_s^k = \bar{\zeta}. \tag{21}$$

Here, ζ is the historical medical expense of patients with a similar condition to the kth insured. The estimated medical expense is the mean value of historical payment.

IV. COEVOLUTIONARY ESTIMATION OF DISTRIBUTION ALGORITHM

The group insurance portfolio problem is so complicated that it is difficult to derive the optimal solution by mathematical calculation. For the single-insured case, an AEDA [7] is proposed to deal with the insurance portfolio problem. However, with the increasing number of insured, the dimension of the problem rises, which makes the optimization problem more complicated. AEDA is effective for the single-insured case while it may not be suitable for this high-dimensional problem. Therefore, this study introduces a coevolutionary approach in collaboration with AEDA and proposes a CEDA.

A. Coevolutionary Method

In general CC method, the objective value, i.e., the approximated expectation payout of the whole group, is utilized to optimize the portfolio plan. However, unlike those inseparable problems, in the proposed group insurance portfolio problem, the approximated payout of each insured can be calculated independently. By using the insured's own payout, optimization can be carried out more efficiently and running resources can be saved. That is to say, the cooperative process is unnecessary.

However, the sum of the premium of the insured should not exceed the total investable amount. Therefore, a coevolutionary process is needed to ensure that the investment amount of each insured is affordable. Another optimization process is executed to optimize the allocation of the total investable amount. All in all, the proposed coevolutionary method is composed of two parts: 1) optimization for investable amount allocation (detail information is given in Section IV-B) and 2) optimization for insurance portfolio of each insured (detail

---

Algorithm 3 Adaptive Estimation of Distribution Algorithm

1: **Initialize** $X[k]$
2: Find the best solution $x[k]\_best$
3: **for** $g = 1$ to G **do**
4:      Find the top NP_{best} solutions
5:      Estimate the probability model
6:      Calculate g_p
7:      Generate a new population
8:      Local Search
9: **end for**

Algorithm 4 Particle Swarm Optimization

1: **Initialize** A
2: **for** $g = 1$ to G_1 **do**
3:      **for** $r = 1$ to NP_1 do
4:          Update $a[r].v$ with(22)
5:          Update $a[r].x$ with(23)
6:          Boundary check
7:      **end for**
8:      **if** $f(a[r].x) > f(a[r].pbest)$
9:          $a[r].pbest = a[r].x$
10:      **if** $f(a[r].x) > f(a\_best)$
11:          $a\_ini = a[r].x$
12:          **end if**
13:      **end if**
14: **end for**

---

information is given in Section IV-C). These two parts are inseparably optimized simultaneously until the termination condition is satisfied. The overall procedure of the coevolutionary method is described as follows.

First, the initial allocation amount is set to be $a\_ini = \{1/n, 1/n, \ldots 1/n\}$ for each insured. Then, the iteration begins until the fitness evaluation time reaches a specified number. a_ini is applied to optimize the portfolio plan for each insured using AEDA given in Algorithm 3. After that, the global fitness value is calculated with each best solution gain after the optimization. If this global fitness value is better than the archived global best fitness value, the global best solution is replaced by this solution and the archived global best allocation amount a_best is replaced by the initial allocation amount a_ini, which is applied in the optimization process in this generation. At the end of the iteration, the allocation amount is optimized by a general PSO given in Algorithm 4 to get the initial allocation amount a_ini for the next generation. Here, the global best solution and global best allocation amount are utilized to execute the optimization process. The detailed information of this process is given in Section IV-B.

The flowchart of the coevolutionary method is given in Fig. 1 and the pseudocode is given in Algorithm 5.

B. Optimization for Investable Amount Allocation

To optimize the allocation of investable amount to each insured, PSO [42], [43], an effective evolutionary algorithm with wide application [44]–[47] is utilized. First, a set of particles with number NP_1 is initialized. The position of each
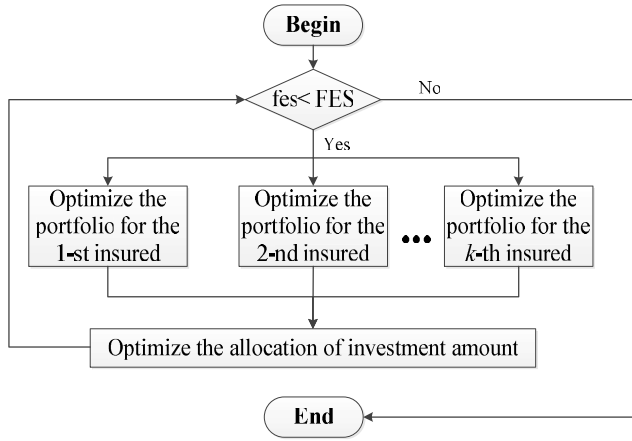
Fig. 1. Flowchart of coevolutionary method.

Algorithm 5 Coevolutionary Method

1: **Initialize** *a_ini*
2: **while** fes < FES **do**
3:    **for** *k* = 1 to *n* **do**
4:        Optimize the investment plan for the *k*-th insured
5:    **end for**
6:    **Calculate** *global_fitness*
7:    **if** *global_fitness* > *global_best_fitness*
8:        *global_best_fitness* = *global_fitness*
9:        *global_gbest* =*g best*
10:        *a_best* = *a_ini*
11:    **end if**
12:    Optimize the allocation of investment amount
13: **end while**
---

particle $a[r].x$ consists of n positive values whose sum is smaller than 1 and the speed of each particle $a[r].v$ set to be 0.

Then, the iteration begins and the update procedure (22) and (23) are applied to each particle in the population

$$a[r].v = \alpha a[r].v + \beta r_1(a[r].\text{pbest} - a[r].x)$$
$$+ \beta r_2(a\_ini - a[r].x) \qquad (22)$$
$$a[r].x = a[r].x + a[r].v \qquad (23)$$

where α is the inertia weight coefficient, β is the learning rate, and r_1 and r_2 are two random real numbers on the interval [0, 1].

After that, a boundary check process is executed to ensure that the basic demand of each insured is guaranteed. Then, the local best solution $a[r].\text{pbest}$ and the global best solution a_ini are updated according to the fitness value. Here, to evaluate the fitness value, the investment amount to each policy is tailored according to the ratio of $a[r].x$ to a_best, i.e.,

$$x = \text{global\_gbest} \times a[r].x / a\_best. \qquad (24)$$

The pseudocode of PSO is given in Algorithm 4.

C. Optimization for Insurance Portfolio of Each Insured

An AEDA, which has been proven to be effective for the single-insured insurance portfolio problem [7], is applied in

this part. With the utilization of the coevolutionary method, the group insurance portfolio problem is decomposed into n single-insured insurance portfolio problem. Each problem can be effectively dealt with using AEDA.

The optimization process begins with the random initialization of the population, in which each solution represents a feasible portfolio plan. To ensure feasibility of solutions, a specific constraint handling method is applied in the initialization process. In this process, the investment amount to each endowment policy is initialized one by one. To ensure that the total premium is affordable for the policyholder, once the premium to a policy is decided, the investable amount to the rest policies should be reduced accordingly. Additionally, the payment period of the policy should also be considered and a corresponding amount should be reduced accordingly. To ensure the randomicity of the initialization process, the purchase order of the policies is upset for each individual.

Then, NP_{best} elite solutions with the largest fitness value in the current population are chosen. A probabilistic model for the investment amount of the endowment policies is estimated by these solutions. The mean value and the variance of these solutions are implemented as the mean value and the variance of the next generation. As for the investment of the hospitalization policy, a histogram method given in (25) is utilized. The solution with better performance has a larger probability to be selected

$$\beta_{i}\_p = \left[\sum_{j=0}^{n_2} \frac{\text{count\_best}_j}{\text{count}_j} \right]^{-1} \frac{\text{count\_best}_i}{\text{count}_i}. \qquad (25)$$

Here, β_{i}_p is the probability for new solutions to choose the ith hospitalization policy; count_i is the number of solutions that choose the ith hospitalization policy in the current population; and count_best_i is the number of solutions that choose the ith hospitalization policy among the elite solutions.

A similar mechanism given by (26) is applied to decide the probabilistic model for the next generation according to the performance of the Gaussian distribution and the Cauchy distribution in the current generation

$$g\_p = \left[\sum_{i=0}^{1} \frac{\text{top}\_d_i}{\text{count}\_d_i} \right]^{-1} \frac{\text{top}\_d_0}{\text{count}\_d_0}. \qquad (26)$$

Here, g_p denotes the probability that samples the next generation by the Gaussian distribution. top_d_i is the number of solutions that sampled from the two distributions among the elite solutions. count_d_i is the total number of solutions that sampled from the two distributions. Here, $i = 0$ denotes the Gaussian distribution while $i = 1$ denotes the Cauchy distribution.

During the generation of the new population, a similar constraint handling method of the initialization process is applied to ensure the feasibility of the solution.

The above processes are executed repeatedly until the termination criterion is satisfied.

TABLE II
SETTINGS OF INCOME

Age	Yearly Income
< 20	12000
$20 \sim 24$	25000
$25 \sim 29$	40000
$30 \sim 34$	50000
$35 \sim 39$	53000
$40 \sim 44$	50000
$45 \sim 49$	48000
$50 \sim 60$	45000
> 60	20000

TABLE III
DETAILED SETTINGS OF THE ENDOWMENT POLICIES

	Payment Period (y_{ij}) and Latest Purchasing Time (q_{ij})					Minimum of the face amount (c_i)
A	5			10		20,000
	75			70		
B	1	5	10	18	25	10,000
	75	65	60	60	55	
C	5			10		10,000
	75			70		
D	6	10	15	20		5,000
	65	60	55	55		
E	1		5	10		1,500
	80		80	75		

TABLE IV
DETAILED SETTINGS OF THE HOSPITALIZATION POLICIES

	Coverage scope	Premium ($\delta_k(t_0)$)
A	critical illness	lowest
B	critical and serious illness	medium
C	critical, serious and general illness	highest

1) AEDA has been proven to be effective for single-insured insurance portfolio problem. To validate the effectiveness of the proposed coevolutionary algorithm, we compare the proposed CEDA with general AEDA without the coevolutionary method.

2) CC method has been proven to be effective for many large-scale optimization problems. To validate that the proposed coevolutionary method is more suitable for the group insurance portfolio problem, we compare the proposed CEDA with the algorithm that combines AEDA with traditional CC method (CCEDA), a contribution-based CC method (CCFR) [48], and a bandit-based CC method (BBCC) [49];

3) CSO [50] algorithm is a latest PSO algorithm that can effectively deal with large-scale optimization problems. jDE [51] algorithm is a classic and representative DE algorithm that is widely applied in many optimization problems. To validate that the proposed algorithm is more suitable for the group insurance portfolio problem than other evolutionary algorithms, we compare the proposed CEDA with CSO and jDE.

For each initial situation, each algorithm runs 30 times and FES is set to 3×10^5. The population size NP of AEDA and CSO is 300 [7]. The population size NP in jDE is 100, while the population size NP_1 for the AEDA component in CEDA and CCEDA is 100 since the problem is decomposed into three subgroups. The population of the PSO component in CEDA is 50. The number of elite solutions selected for model estimation NP_{best} is 45% of the whole population. The control parameter φ in CSO is 0.1 [50]. In the PSO component in CEDA, α is 0.729 and β is 1.49445 [52].

V. EXPERIMENTAL RESULTS

A. Experiment Settings

We consider the situations that the group consists of three insured with age {30, 35, 45} simulating a company or {5, 30, 35} simulating a family. The initial investable amount I_1 of the whole group varying in the range of {200000, 500000, 1000000} and the duration varying in the range of {10, 30}. The yearly income of each person, which is set according to [7], is supposed to be the same and showed in Table II. The mortality rate of each insured and the incidence rate of diseases are set in accordance with the regular patterns. The insurance policies considered in this study are tabulated in Tables III and IV [7]. The parameters of the insurance policies are set according to real life insurance products.

To evaluate the effectiveness of the proposed portfolio investment strategy, we compare the performance of the proposed strategy with the other two strategies.

1) To validate the effectiveness of the proposed portfolio strategy and explain the superiority of portfolio strategy, we compare the portfolio strategy with the single-policy investment strategy, which chooses the endowment policy with the highest payout (SP).

2) To validate the necessity of the asset allocation component in CEDA, we compare the proposed CEDA with CEDA that allocates the entire investable asset averagely (CEDA-a).

To evaluate the effectiveness of the proposed CEDA, we compare the performance of the proposed CEDA with other six algorithms.

B. Parameter Analysis

The parameters in Section V-A are all carried forward from the existing literature [7], [50]–[52]. However, as for the iteration times of the AEDA component and PSO component in CEDA, since the coevolutionary method is newly proposed in this article, the setting of these two parameters has not been involved in any existing literature. Experiments are conducted to explore the characteristics of these two parameters. Under the premise of ensuring that the total fitness evaluation times are set to be 3×10^5, the iteration times of both of these two components are set to be 20, 50, 100, and 200. Experiments are made in the following four situations.

1) $t_0 = 30, 35, 45$; $T_3 = 10$; $I_1 = 200000$.
2) $t_0 = 30, 35, 45$; $T_3 = 30$; $I_1 = 500000$.
3) $t_0 = 5, 30, 35$; $T_3 = 10$; $I_1 = 500000$.
4) $t_0 = 5, 30, 35$; $T_3 = 30$; $I_1 = 1000000$.

A surface graph is a figure that can reflect the relationship of the dependent variable and two independent variables. It
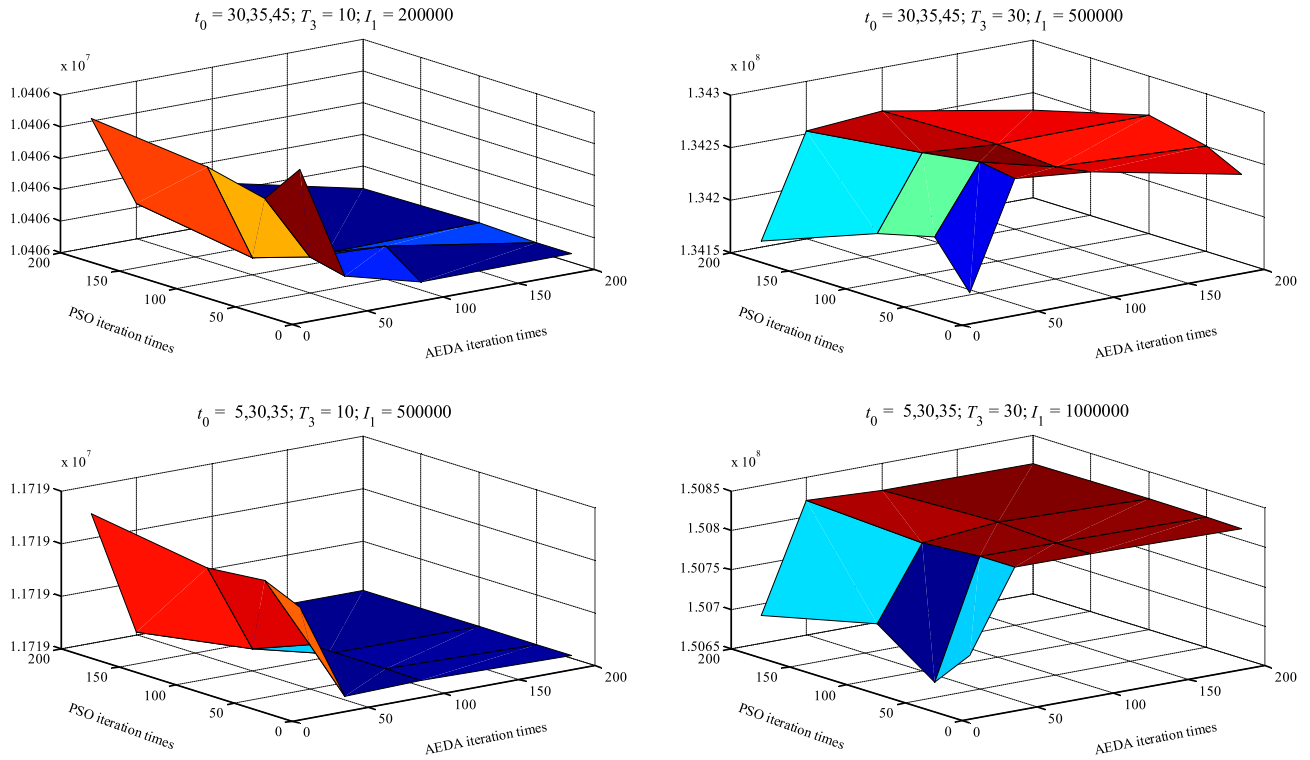
Fig. 2.   Surface graph for parameter analysis.

is widely applied to find the best combination between two sets of data [53]. In this study, surface graphs that reflect the impact of PSO iteration times and AEDA iteration times on the fitness value are utilized to decide the setting of these two parameters. The results are shown in Fig. 2. From this figure, it can be seen that the iteration times of the PSO component does not have a significant impact on the performance of the algorithm. It may because that the iteration times of the PSO component do not have an obvious impact of the optimization results. As for iteration times of the AEDA component, in situations 1) and 3), the algorithm performs worse and worse with the increase of the iteration times. In situations 2) and 4), the algorithm performs better when the generation time is relatively higher. In addition, it can be easily found out that the surface graph of situations 1) and 3) is similar while the surface graph of situations 2) and 4) is similar. From this phenomenon, we can make a conjecture that the performance of these two parameters is related to the duration considered in the problem. This may because that the characteristic of the optimization problem changes significantly as the duration changes. Since the payment period of some policies is longer than ten years, investing in these policies is not a suitable choice when the investment duration is ten years. Nevertheless, when the investment duration extends to thirty years, these policies may perform well.

It can be discovered that the best setting of the iteration times of PSO component and AEDA component is different with different situations. However, in situations 1) and 3), although the performance becomes worse and worse with the increase of iteration times of the AEDA component, the difference is relatively small. Therefore, the iteration times

of the AEDA component are set to be 100, whose performances in situations 2) and 4) are good and in situations 1) and 3) are also acceptable. The iteration times of PSO are set to be 50, whose performances are acceptable in all situations.

C. Comparison of the Investment Strategies

The experimental results obtained in 30 independent runs of each algorithm are given in Table V for comparison. The result of the Wilcoxon rank sum test [54] is applied to examine whether the difference between the algorithms is significant. A *p*-value smaller than 0.05 indicates a significant difference between two algorithms. The smaller the *p*-value is, the more significant the difference is.

1) Comparison Between Portfolio Investment Strategy and Single-Policy Investment Strategy: To validate the rationality of the portfolio investment strategy, comparisons are made between the proposed CEDA with an SP strategy. Table V shows that the portfolio investment strategy outperforms the SP investment strategy in all of the 12 situations. In two of the situations, the *p*-value is 10^{-14} orders of magnitude; in four of the situations, the *p*-value is 10^{-13} orders of magnitude while in the other six, the *p*-value is 10^{-11} orders of magnitude. These data show that the portfolio strategy outperforms the SP strategy significantly, which means that the portfolio strategy is more effective than single-policy strategy that just invests in one insurance policy. Actually, few people invest in just one kind of insurance policy in real life. As a famous proverb goes, it is unwise to put all your eggs in one basket. None of an insurance policy performs better than another policy at any time for any insured. Otherwise, the worse one will

TABLE V
COMPARISON OF CEDA WITH OTHER ALGORITHMS

		CEDA	SP	CEDA-a	AEDA	CCEDA	CCFR	CCBB	CSO	jDE
$t_0 = 30,35,45$ $T_3 = 10$ $I_1 = 200000$	Max	1.041E+07	8.430E+06	1.041E+07	1.041E+07	1.039E+07	1.040E+07	1.040E+07	1.008E+07	1.027E+07
	Mean	1.041E+07	8.430E+06	1.041E+07	1.040E+07	1.032E+07	1.033E+07	1.033E+07	9.494E+06	1.018E+07
	Std	6.005E+00	0.000E+00	5.029E+01	2.145E+04	7.605E+04	7.366E+04	6.712E+04	2.878E+05	4.713E+04
	p-value	-	2.708E-14$^+$	9.485E-14$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$
$t_0 = 30,35,45$ $T_3 = 10$ $I_1 = 500000$	Max	1.364E+07	1.105E+07	1.364E+07	1.364E+07	1.364E+07	1.364E+07	1.363E+07	1.322E+07	1.861E+07
	Mean	1.364E+07	1.091E+07	1.364E+07	1.363E+07	1.357E+07	1.352E+07	1.349E+07	1.237E+07	1.821E+07
	Std	9.472E-09	5.471E+05	5.026E+01	4.485E+04	7.273E+04	1.124E+05	2.078E+05	4.824E+05	1.455E+05
	p-value	-	8.713E-14$^+$	6.140E-14$^+$	1.212E-12$^+$	1.212E-12$^+$	1.212E-12$^+$	1.212E-12$^+$	1.212E-12$^+$	3.160E-12$^+$
$t_0 = 30,35,45$ $T_3 = 10$ $I_1 = 1000000$	Max	1.904E+07	1.543E+07	1.903E+07	1.904E+07	1.902E+07	1.903E+07	1.903E+07	1.898E+07	1.102E+08
	Mean	1.904E+07	1.517E+07	1.903E+07	1.902E+07	1.883E+07	1.879E+07	1.888E+07	1.757E+07	1.077E+08
	Std	6.244E+00	9.636E+05	5.032E+01	7.135E+04	1.599E+05	1.478E+05	1.273E+05	6.884E+05	1.184E+06
	p-value	-	3.721E-13$^+$	9.485E-14$^+$	3.160E-12$^+$	3.160E-12$^+$	3.160E-12$^+$	3.160E-12$^+$	3.160E-12$^+$	3.020E-11$^+$
$t_0 = 30,35,45$ $T_3 = 30$ $I_1 = 200000$	Max	1.135E+08	9.154E+07	1.132E+08	1.132E+08	1.134E+08	1.130E+08	1.134E+08	1.096E+08	1.313E+08
	Mean	1.134E+08	9.153E+07	1.132E+08	1.114E+08	1.082E+08	9.284E+07	1.070E+08	1.045E+08	1.280E+08
	Std	6.291E+04	9.015E+03	4.209E+04	3.599E+06	3.837E+06	3.718E+07	4.055E+06	2.767E+06	1.319E+06
	p-value	-	1.136E-11$^+$	3.020E-11$^+$	3.020E-11$^+$	6.066E-11$^+$	3.001E-11$^+$	4.975E-11$^+$	3.020E-11$^+$	3.020E-11$^+$
$t_0 = 30,35,45$ $T_3 = 30$ $I_1 = 500000$	Max	1.344E+08	1.056E+08	1.339E+08	1.343E+08	1.341E+08	1.340E+08	1.340E+08	1.294E+08	1.644E+08
	Mean	1.343E+08	1.056E+08	1.338E+08	1.336E+08	1.285E+08	1.278E+08	1.277E+08	1.249E+08	1.615E+08
	Std	2.780E+04	9.369E+03	3.086E+04	4.506E+05	4.966E+06	4.131E+06	4.145E+06	2.237E+06	1.356E+06
	p-value	-	1.337E-11$^+$	3.020E-11$^+$	1.464E-10$^+$	3.020E-11$^+$	3.020E-11$^+$	3.020E-11$^+$	3.020E-11$^+$	3.020E-11$^+$
$t_0 = 30,35,45$ $T_3 = 30$ $I_1 = 1000000$	Max	1.688E+08	1.291E+08	1.683E+08	1.688E+08	1.685E+08	1.682E+08	1.686E+08	1.648E+08	8.375E+06
	Mean	1.688E+08	1.291E+08	1.682E+08	1.680E+08	1.637E+08	1.567E+08	1.624E+08	1.577E+08	8.195E+06
	Std	1.643E+04	4.281E+05	4.520E+04	7.294E+05	5.159E+06	3.000E+07	3.372E+06	2.771E+06	7.210E+04
	p-value	-	1.539E-11$^+$	3.020E-11$^+$	5.573E-10$^+$	3.020E-11$^+$	3.020E-11$^+$	3.020E-11$^+$	3.020E-11$^+$	1.720E-12$^+$
$t_0 = 5,30,35$ $T_3 = 10$ $I_1 = 200000$	Max	8.478E+06	6.868E+06	8.477E+06	8.478E+06	8.477E+06	8.476E+06	8.477E+06	8.223E+06	1.129E+07
	Mean	8.478E+06	6.723E+06	8.477E+06	8.470E+06	8.383E+06	8.411E+06	8.404E+06	7.727E+06	1.119E+07
	Std	3.657E-01	3.277E+05	1.778E+01	1.951E+04	1.464E+05	7.409E+04	5.258E+04	2.114E+05	6.024E+04
	p-value	-	1.784E-13$^+$	4.289E-14$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$
$t_0 = 5,30,35$ $T_3 = 10$ $I_1 = 500000$	Max	1.172E+07	9.493E+06	1.172E+07	1.172E+07	1.172E+07	1.171E+07	1.172E+07	1.150E+07	1.655E+07
	Mean	1.172E+07	9.420E+06	1.172E+07	1.171E+07	1.154E+07	1.162E+07	1.160E+07	1.067E+07	1.619E+07
	Std	6.090E-01	3.966E+05	1.789E+01	3.570E+04	2.744E+05	7.148E+04	1.145E+05	4.156E+05	1.339E+05
	p-value	-	5.312E-13$^+$	6.505E-14$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$	1.720E-12$^+$	1.212E-12$^+$
$t_0 = 5,30,35$ $T_3 = 10$ $I_1 = 1000000$	Max	1.712E+07	1.387E+07	1.711E+07	1.712E+07	1.711E+07	1.711E+07	1.711E+07	1.676E+07	9.372E+07
	Mean	1.712E+07	1.363E+07	1.711E+07	1.709E+07	1.699E+07	1.701E+07	1.694E+07	1.565E+07	9.174E+07
	Std	7.578E-09	8.942E+05	1.778E+01	8.816E+04	1.388E+05	1.056E+05	1.587E+05	6.300E+05	8.552E+05
	p-value	-	4.151E-13$^+$	2.708E-14$^+$	1.212E-12$^+$	1.212E-12$^+$	1.212E-12$^+$	1.212E-12$^+$	1.212E-12$^+$	3.020E-11$^+$
$t_0 = 5,30,35$ $T_3 = 30$ $I_1 = 200000$	Max	9.534E+07	7.654E+07	9.523E+07	9.511E+07	9.525E+07	9.487E+07	9.527E+07	9.358E+07	1.133E+08
	Mean	9.531E+07	7.654E+07	9.523E+07	9.350E+07	9.037E+07	8.405E+07	9.119E+07	8.906E+07	1.110E+08
	Std	3.844E+04	5.352E+03	1.190E+04	2.970E+06	2.967E+06	2.312E+07	3.188E+06	2.104E+06	9.985E+05
	p-value	-	1.015E-11$^+$	5.573E-10$^+$	3.020E-11$^+$	3.690E-11$^+$	3.018E-11$^+$	4.077E-11$^+$	3.020E-11$^+$	3.020E-11$^+$
$t_0 = 5,30,35$ $T_3 = 30$ $I_1 = 500000$	Max	1.162E+08	9.069E+07	1.160E+08	1.161E+08	1.160E+08	1.160E+08	1.161E+08	1.126E+08	1.473E+08
	Mean	1.162E+08	9.068E+07	1.160E+08	1.152E+08	1.123E+08	9.964E+07	1.122E+08	1.094E+08	1.451E+08
	Std	1.482E+04	5.826E+03	1.471E+04	2.136E+06	3.238E+06	3.396E+07	3.315E+06	1.978E+06	1.344E+06
	p-value	-	1.449E-11$^+$	3.020E-11$^+$	3.338E-11$^+$	3.020E-11$^+$	3.012E-11$^+$	3.020E-11$^+$	3.020E-11$^+$	3.020E-11$^+$
$t_0 = 5,30,35$ $T_3 = 30$ $I_1 = 1000000$	Max	1.509E+08	1.143E+08	1.506E+08	1.508E+08	1.503E+08	1.504E+08	1.507E+08	1.503E+08	1.861E+07
	Mean	1.508E+08	1.143E+08	1.506E+08	1.493E+08	1.446E+08	1.303E+08	1.468E+08	1.432E+08	1.821E+07
	Std	1.477E+04	7.355E+03	1.427E+04	3.859E+06	4.503E+06	4.440E+07	4.009E+06	3.446E+06	1.455E+05
	p-value	-	1.337E-11$^+$	3.020E-11$^+$	3.690E-11$^+$	3.020E-11$^+$	3.012E-11$^+$	3.020E-11$^+$	3.020E-11$^+$	3.160E-12$^+$

not exist in the insurance market. The payout can be optimized to the greatest extent through the portfolio strategy since some policy may perform well at some time while the others may perform well at other times.

2) Comparison Between Heterogeneous Allocation Strategy and Average Allocation Strategy: To validate the rationality of heterogeneous allocation strategy, comparisons are made between the proposed CEDA with CEDA-a. The proposed

TABLE VI
ASSET ALLOCATION IN EACH SITUATION

	A	B	C
$t_0 = 30,35,45$; $T_3 = 10$; $I_1 = 200000$	159457	19932	19932
$t_0 = 30,35,45$; $T_3 = 10$; $I_1 = 500000$	399457	49932	49932
$t_0 = 30,35,45$; $T_3 = 10$; $I_1 = 1000000$	799457	99932	99932
$t_0 = 30,35,45$; $T_3 = 30$; $I_1 = 200000$	1228208	153526	150335
$t_0 = 30,35,45$; $T_3 = 30$; $I_1 = 500000$	1464172	183524	183526
$t_0 = 30,35,45$; $T_3 = 30$; $I_1 = 1000000$	1868208	233474	233495
$t_0 = 5,30,35,$; $T_3 = 10$; $I_1 = 200000$	159797	19974	19974
$t_0 = 5,30,35,$; $T_3 = 10$; $I_1 = 500000$	399797	49974	49974
$t_0 = 5,30,35,$; $T_3 = 10$; $I_1 = 1000000$	799797	99974	99974
$t_0 = 5,30,35,$; $T_3 = 30$; $I_1 = 200000$	984525	121845	121845
$t_0 = 5,30,35,$; $T_3 = 30$; $I_1 = 500000$	1224714	153089	152319
$t_0 = 5,30,35,$; $T_3 = 30$; $I_1 = 1000000$	1624637	203089	203089

CEDA outperforms CEDA-a in all of the 12 situations. In six of the situations, the *p*-value is 10^{-14} orders of magnitude; in five of the situations, the *p*-value is 10^{-11} orders of magnitude while in the other one, the *p*-value is 10^{-10} orders of magnitude. From these results, it can be found out that by properly allocating the amount of investment for each insured, the payout can be maximized. When determining insurance portfolio plan, each insured are of different status in the group. The asset allocation of the best portfolio plan in 30 runs optimized by CEDA is given in Table VI. From this table, it can be discovered that with the change of the initial wealth or duration of investment, the asset allocation should be adjusted accordingly.

D. Comparison of the Algorithms

1) Comparison Between Collaborative Optimization Strategy and Overall Optimization Strategy: In the optimization of single-insured insurance portfolio problem, an AEDA algorithm has been proven effective [7]. To verify whether this algorithm is also suitable for the group cases, comparisons are made between the proposed CEDA and AEDA. Table V shows that the proposed CEDA outperforms AEDA in all of the 12 situations. In six of the situations, the *p*-value is 10^{-12} orders of magnitude; in four of the situations, the *p*-value is 10^{-11} orders of magnitude while in the other two, the *p*-value is 10^{-10} orders of magnitude. These results indicate that CEDA is more suitable for group portfolio problem than AEDA. With the increase of insured considered in the portfolio problem, the possible combination of the purchase of endowment policies increases exponentially. The dimension of the optimization problem also increases exponentially. Since AEDA is not effective when dealing with high-dimensional problems, it performs worse than the proposed CEDA in the group insurance portfolio problem.

2) Comparison Between Different Collaborative Optimization Strategy: The traditional CC method [12] has also been proven to be effective to deal with high dimensional problems. To further validate the effectiveness of the proposed coevolutionary method, comparisons are made between the proposed CEDA with CCEDA, CCFR, and BBCC. Table V shows that the proposed CEDA outperforms the other three algorithms in all of the 12 situations. In six of the situations, the *p*-value is 10^{-12} orders of magnitude while

in the other six, the *p*-value is 10^{-11} orders of magnitude. These results show that the proposed coevolutionary method is more suitable for the group insurance portfolio problem than other CC methods. In the other three CC methods, each subcomponent is separated completely and evolves independently. However, in the proposed group insurance portfolio problem, the constraint of total investable assets connects all of the insured in the group together and makes the problem inseparable. Therefore, the proposed coevolutionary method, which takes the interaction between subcomponents into consideration, is more suitable for the proposed inseparable group insurance portfolio problem.

3) Comparison With Other Evolutionary Algorithms: To further validate the effectiveness of the proposed algorithm, comparisons are made between the proposed algorithm and two other evolutionary algorithms, CSO and jDE. Table V shows that the proposed CEDA outperforms CSO and jDE in all of the 12 stations. In six of the situations, the *p*-value is 10^{-12} orders of magnitude while in the other six; the *p*-value is 10^{-11} orders of magnitude for both of these two algorithms. These results show that the proposed CEDA is more suitable for the group insurance portfolio problem than other evolutionary algorithms. This is because the group insurance portfolio problem has both of continuous and discrete decision variables. The discrete decision variables β_k, which denotes the types of hospitalization policies to be chosen, are categorical but not ordinal. It is inefficient to deal with the considered problem by using continuous evolutionary algorithms, such as PSO and DE with the rounding method. Therefore, we choose EDA as an optimization algorithm due to its great ability to deal with both continuous and discrete variables.

E. Comparison of Convergence Speed

The convergence figures of the above algorithms besides SP strategy, CCFR, and BBCC in the following four situations are given in Fig. 3.
1) $t_0 = 30, 35, 45$; $T_3 = 10$; $I_1 = 200000$.
2) $t_0 = 30, 35, 45$; $T_3 = 30$; $I_1 = 500000$.
3) $t_0 = 5, 30, 35$; $T_3 = 10$; $I_1 = 500000$.
4) $t_0 = 5, 30, 35$; $T_3 = 30$; $I_1 = 1000000$.

Since the SP strategy just needs to figure out a policy with the highest payout, this optimization problem is much simpler than the portfolio problem. It makes no sense to compare this algorithm with the other ones. As for CCFR and BBCC, the convergence speeds of them are quite similar to CCEDA. To avoid mess in the figure, we take CCEDA as the representation of CC framework-based algorithms.

Since the only difference between CEDA-a and CEDA is the subtraction of asset allocation strategy, the convergence trends of these two algorithms are similar. However, in CEDA-a, the asset allocation has already been determined, so that the convergence speed is much faster than CEDA. But the performances of CEDA are always better than CEDA-a due to the effectiveness of the asset allocation strategy. CCEDA, CSO, and jDE converge fast in the early stage of optimization but the final optimization results are not as good as the other algorithms. In contrast, AEDA converges slow and performs
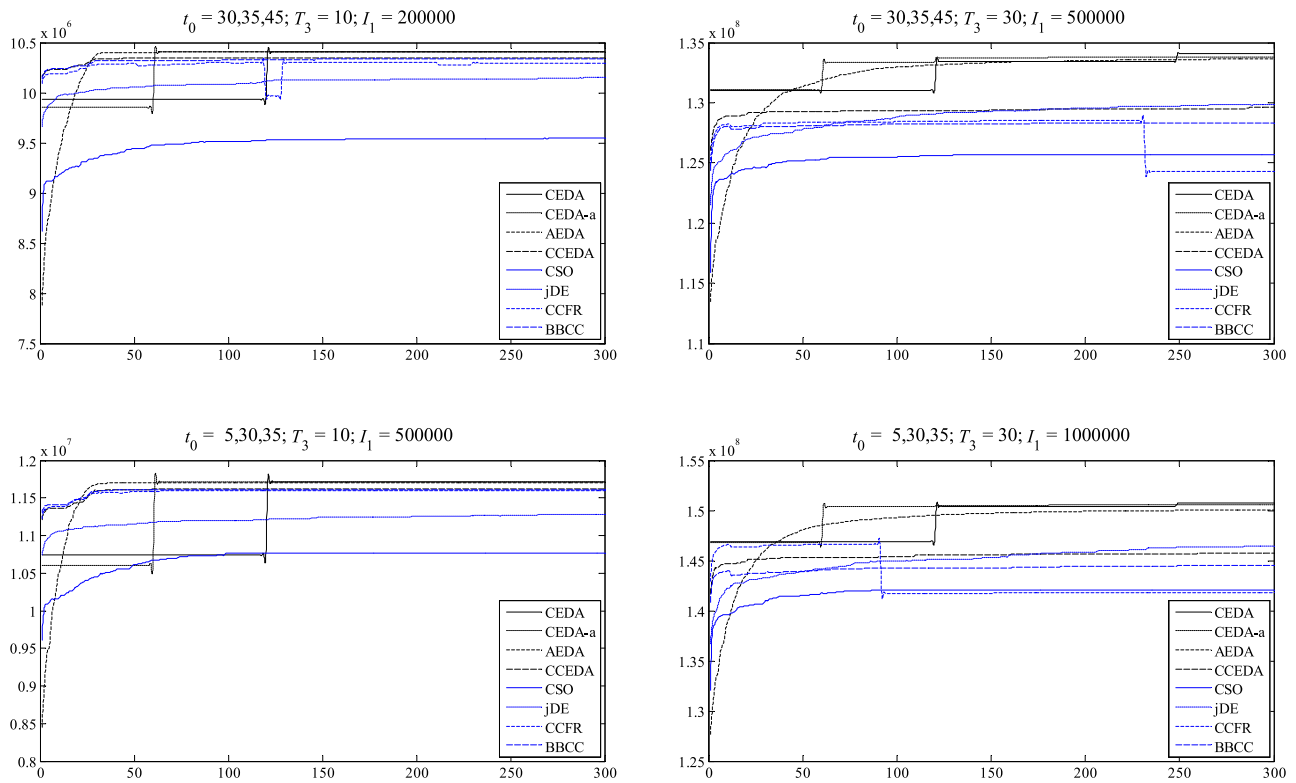
Fig. 3. Convergence curves of each algorithm.

the worst in the early stage of optimization, but has the ability to convergence to better solutions than CCEDA, CSO, and jDE. In general, the proposed CEDA not only converges fast but also can find the best solution in all of the four situations.

F. Stochastic Analyzes

Since the model applied in the experiments is an approximation model, the suitability of this model for the group insurance portfolio problem should be analyzed. Due to space limitations, this part is given in the supplementary material.

VI. CONCLUSION

In this article, a data-driven model is proposed to approximate the expectation payout of the group insurance portfolio problem. The objective of the problem is to maximize the total payout with fixed investable amount. As the number of insured increases, the combination of insurance portfolio increases exponentially, which increases the dimensionality of the problem as well.

In response to the increase in dimensions, a new coevolutionary EDA is devised. The coevolutionary method proposed in this article decomposes the group portfolio problem into several single-insured portfolio problems. A classic PSO algorithm is utilized to optimize the asset allocation to each insured and an AEDA we proposed previously is utilized to optimize the portfolio problem for each insured.

This article explores the use of evolutionary algorithms in the optimization of the insurance portfolio problem for a group. However, in this article, the problem has been simplified and many constraints have been ignored. For example,

people with certain characteristic cannot invest in certain insurance policies in practical. In subsequent research, those constraints of insurance purchase should be taken into account. What is more, this article studies instantaneous insurance portfolio problem, which can extend to the long-term insurance portfolio problem in future researches.

REFERENCES

[1] M. E. Yaari, "Uncertain lifetime, life insurance, and the theory of the Consumer," *Rev. Econ. Stud.*, vol. 32, no. 2, pp. 137–150, 1965.

[2] N. H. Hakansson, "Optimal investment and consumption strategies under risk, an uncertain lifetime, and insurance," *Int. Econ. Rev.*, vol. 10, no. 3, pp. 443–466, 1969.

[3] S. F. Richard, "Optimal consumption, portfolio and life insurance rules for an uncertain lived individual in a continuous time model," *J. Financ. Econ.*, vol. 2, no. 2, pp. 187–203, 1975.

[4] S.R. Pliska and J. Ye, "Optimal life insurance purchase and consumption/investment under uncertain lifetime," *J. Bank. Finan.*, vol. 31, no. 5, pp. 1307–1319, 2007.

[5] J. Ye, "Optimal life insurance, consumption and portfolio under uncertainty: Martingale methods," in *Proc. Amer. Control Conf.*, 2007, pp. 1103–1109.

[6] A. S. Mousa, D. Pinheiro, and A. A. Pinto, "Optimal life-insurance selection and purchase within a market of several life-insurance providers," *Insur. Math. Econ.*, vol. 67, pp. 133–141, Mar. 2016.

[7] W. Shi, W. N. Chen, Y. Lin, T. Gu, S. Kwong, and J. Zhang, "An adaptive estimation of distribution algorithm for multi-policy insurance investment planning," *IEEE Trans. Evol. Comput.*, vol. 23, no. 1, pp. 1–14, Feb. 2019.

[8] H. Huang, M. A. Milevsky, and J. Wang, "Portfolio choice and life insurance: The CRRA case," *J. Risk Insur.*, vol. 75, no. 4, pp. 847–872, 2008.

[9] M. Kwak, Y. H. Shin, and U. J. Choi, "Optimal investment and consumption decision of family with life insurance," *Insur. Math. Econ.*, vol. 48, no. 2, pp. 176–188, 2011.

[10] K. Bruhn and M. Steffensen, "Household consumption, investment and life insurance," *Insur. Math. Econ.*, vol. 48, no. 3, pp. 315–325, 2011.

[11] E. Bengoetxea, P. Larrañaga, I. Bloch, and A. Perchant, "Estimation of distribution algorithms," *Genet. Algorithms Evol. Comput.*, vol. 64, no. 5, pp. 1140–1148, 2001.

[12] M. A. Potter and K. A. De Jong, "A cooperative coevolutionary approach to function optimization," in *Parallel Problem Solving from Nature—PPSN III*. Jerusalem, Israel: Springer, 1994 pp. 249–257.

[13] I. Duarte, D. Pinheiro, A. A. Pinto, and S. R. Pliska, "An overview of optimal life insurance purchase, consumption and investment problems," in *Dynamics, Games and Science I*. Braga, Portugal: Springer, 2011, pp. 271–286.

[14] I. Duarte, A. A. Pinto, and S. R. Pliska, "Optimal life insurance purchase, consumption and investment on a financial market with multi-dimensional diffusive terms," *Optimzation*, vol. 63, no. 11, pp. 1737–1760, 2014.

[15] G. Stabile, "Optimal timing of the annuity purchase: Combined stochastic control and optimal stopping problem," *Int. J. Theor. Appl. Finan.*, vol. 9, no. 2, pp. 151–170, 2006.

[16] T.A. Pirvu and H. Zhang, "Optimal investment, consumption and life insurance under mean-reverting returns: The complete market solution," *Insur. Math. Econ.*, vol. 51, no. 2, pp. 303–309, 2012.

[17] M.T. Kronborg and M. Steffensen, "Optimal consumption, investment and life insurance with surrender option guarantee," *Scand. Actuar. J.*, vol. 2015, no. 1, pp. 59–87, 2015.

[18] Y. Shen and J. Wei, "Optimal investment-consumption-insurance with random parameters," *Scand. Actuar. J.*, vol. 2016, no. 1, pp. 37–62, 2016.

[19] M. S. Dorfman, *Introduction to Risk Management and Insurance*. Upper Saddle River, NJ, USA: Pearson Prentice Hall, 2004.

[20] M. Gong, H. Li, E. Luo, J. Liu, and J. Liu, "A multiobjective cooperative coevolutionary algorithm for hyperspectral sparse unmixing," *IEEE Trans. Evol. Comput.*, vol. 21, no. 2, pp. 234–248, Apr. 2017.

[21] N. R. Sabar, J. Abawajy, and J. Yearwood, "Heterogeneous cooperative co-evolution memetic differential evolution algorithm for big data optimization problems," *IEEE Trans. Evol. Comput.*, vol. 21, no. 2, pp. 315–327, Apr. 2017.

[22] M. H. Nguyen, H. A. Abbass, and R. I. Mckay, "Analysis of CCME: Coevolutionary dynamics, automatic problem decomposition, and regularization," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 38, no. 1, pp. 100–109, Jan. 2008.

[23] J. Derrac, I. Triguero, S. Garcia, and F. Herrera, "Integrating instance selection, instance weighting, and feature weighting for nearest neighbor classifiers by coevolutionary algorithms," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 5, pp. 1383–1397, Oct. 2012.

[24] Y. Liu, X. Yao, Q. Zhao, and T. Higuchi, "Scaling up fast evolutionary programming with cooperative coevolution," in *Proc. Congr. Evol. Comput.*, 2001, pp. 1101–1108.

[25] Y. J. Shi, H. F. Teng, and Z. Q. Li, "Cooperative co-evolutionary differential evolution for function optimization," in *Advances in Natural Computation*. Changsha, China: Springer, 2005, pp. 1080–1088.

[26] F. van den Bergh and A. P. Engelbrecht, "A cooperative approach to particle swarm optimization," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 225–239, Jun. 2004.

[27] M. N. Omidvar, X. Li, Y. Mei, and X. Yao, "Cooperative co-evolution with differential grouping for large scale optimization," *IEEE Trans. Evol. Comput.*, vol. 18, no.3, pp. 378–393, Jun. 2014.

[28] Y. Sun, M. Kirley, and S. K. Halgamuge, "Extended differential grouping for large scale global optimization with direct and indirect variable interactions," in *Proc. Genet. Evol. Comput. Conf.*, 2015, pp. 313–320.

[29] Y. Mei, X. Yao, X. Li, and M. N. Omidvar, "A competitive divide-and-conquer algorithm for unconstrained large scale black-box optimization," *ACM Trans. Math. Softw.*, vol. 42, no. 3, p. 13, 2016.

[30] Z. Yang, K. Tang, and X. Yao, "Multilevel cooperative coevolution for large scale optimization," in *Proc. Congr. Evol. Comput.*, 2008, pp. 1663–1670.

[31] M. N. Omidvar, X. Li, and X. Yao, "Cooperative co-evolution with delta grouping for large scale non-separable function optimization," in *Proc. Congr. Evol. Comput.*, 2010, pp. 1–8.

[32] J. Holland, "Genetic algorithms," *Sci. Amer.*, vol. 267, no. 1, pp. 66–72, 1992.

[33] Z. Yang, K. Tang, and X. Yao, "Differential evolution for high-dimensional function optimization," in *Proc. Congr. Evol. Comput.*, 2007, pp. 3523–3530.

[34] T. Ray and X. Yao, "A cooperative coevolutionary algorithm with correlation based adaptive variable partitioning," in *Proc. Congr. Evol. Comput*, 2009, pp. 983–989.

[35] M. N. Omidvar, X. Li, Z. Yang, and X. Yao, "Cooperative co-evolution for large scale optimization through more frequent random grouping," in *Proc. Congr. Evol. Comput.*, 2010, pp. 1–8.

[36] R. Storn and K. Price, "Differential evolution—A simple and efficient heuristic for global optimization over continuous spaces," *J. Global Optim.*, vol. 11, no. 4, pp. 341–359, 1997.

[37] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. IEEE Int. Conf. Neural Netw.*, 1995, pp. 1942–1948.

[38] Y. Chen, X. Sun, D. Gong, Y. Zhang, J. Choi, and S. Klasky, "Personalized search inspired fast interactive estimation of distribution algorithm and its application," *IEEE Trans. Evol. Comput.*, vol. 21, no. 4, pp. 588–600, Aug. 2017.

[39] S. Y. Wang and L. Wang, "An estimation of distribution algorithm-based memetic algorithm for the distributed assembly permutation flow-shop scheduling problem," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 46, no. 1, pp. 139–149, Jan. 2016.

[40] Q. Lu and X. Yao "Clustering and learning Gaussian distribution for continuous optimization," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 35, no. 2, pp. 195–204, May 2005.

[41] V. A. Shim, K. C. Tan, and C. Y. Cheong, "A hybrid estimation of distribution algorithm with decomposition for solving the multiobjective multiple traveling salesman problem," *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 5, pp. 682–691, Sep. 2012.

[42] S. Zhang, J. Xu, L. H. Lee, E. P. Chew, W. P. Wong, and C.-H. Chen, "Optimal computing budget allocation for particle swarm optimization in stochastic optimization," *IEEE Trans. Evol. Comput.*, vol. 21, no. 2, pp. 206–219, Apr. 2017.

[43] Q. Lin *et al.*, "Particle swarm optimization with a balanceable fitness estimation for many-objective optimization problems," *IEEE Trans. Evol. Comput.*, vol. 22, no. 1, pp. 32–46, Feb. 2018.

[44] W. Dong and M. C. Zhou, "A supervised learning and control method to improve particle swarm optimization algorithms," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 7, pp. 1135–1148, Jul. 2017.

[45] Y.-H. Jia *et al.*, "A dynamic logistic dispatching system with set-based particle swarm optimization," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 48, no. 9, pp. 1607–1621, Sep. 2018.

[46] W. Yuan, Y. Liu, H. Wang, and Y. Cao, "A geometric structure-based particle swarm optimization algorithm for multiobjective problems," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 9, pp. 2516–2537, Sep. 2017.

[47] J.-T. Tsai, P.-Y. Chou, and J.-H. Chou, "Color filter polishing optimization using ANFIS with sliding-level particle swarm optimizer," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 3, pp. 1193–1207, Mar. 2020.

[48] M. Yang *et al.*, "Efficient resource allocation in cooperative co-evolution for large-scale global optimization," *IEEE Trans. Evol. Comput.*, vol. 21, no. 4, pp. 493–505, Aug. 2017.

[49] B. Kazimipour, M. N. Omidvar, A. K. Qin, X. Li, and X. Yao, "Bandit-based cooperative coevolution for tackling contribution imbalance in large-scale optimization problems," *Appl. Soft Comput.*, vol. 76, pp. 265–281, Mar. 2019.

[50] R. Cheng and Y. Jin, "A competitive swarm optimizer for large scale optimization," *IEEE Trans. Cybern.*, vol. 45, no. 2, pp. 191–204, Feb. 2015.

[51] J. Brest, S. Greiner, B. Boskovic, V. Zumer, and M. Mernik, "Self-adapting control parameters in differential evolution: A comparative study on numerical benchmark problems," *IEEE Trans. Evol. Comput.*, vol. 10, no. 6, pp. 646–657, Dec. 2006.

[52] R. C. Eberhart and Y. Shi, "Comparing inertia weights and constriction factors in particle swarm optimization," in *Proc. Congr. Evol. Comput.*, 2000, pp. 84–88.

[53] A. M. A. Lorza *et al.*, "Carotid artery lumen segmentation in 3D free-hand ultrasound images using surface graph cuts,," *Medical Image Computing and Comput-Assisted Invtervention (MICCAI)*. Nagoya, Japan: Springer, 2013, pp. 542–549.

[54] F. Wilcoxon, S. K. Katti, and R. A. Wilcox, "Critical values and probability levels for the Wilcoxon rank sum test and the Wilcoxon signed rank test," in *Selected Tables in Mathematical Statistics*, vol. 1. Providence, RI, USA Amer. Math. Soc., 1970, pp. 171–259.

Wen Shi (Student Member, IEEE) received the bachelor's degree in applied mathematics from Sun Yat-sen University, Guangzhou, China, in 2016. She is currently pursuing the Ph.D. degree with the School of Computer Science and Engineering, South China University of Technology, Guangzhou.

Her current research interests include evolutionary computation algorithms and their applications on financial problems. So far, she specifically works on differential evolution, particle swarm optimization, and estimation of distribution algorithms and their applications on financial problems, such as insurance investment optimization problem.

Wei-Neng Chen (Senior Member, IEEE) received the bachelor's and Ph.D. degrees in computer science from Sun Yat-sen University, Guangzhou, China, in 2006 and 2012, respectively.

Since 2016, he has been a Full Professor with the School of Computer Science and Engineering, South China University of Technology, Guangzhou, where he is currently the Vice Dean of the School. He has coauthored over 100 international journal and conference papers, including more than 50 papers published in the IEEE transactions journals. His current research interests include computational intelligence, swarm intelligence, network science, and their applications.

Prof. Chen was a recipient of the IEEE Computational Intelligence Society Outstanding Dissertation Award in 2016, and the National Science Fund for Excellent Young Scholars in 2016. He is currently the Vice-Chair of the IEEE Guangzhou Section. He is also a Committee Member of the IEEE CIS Emerging Topics Task Force. He serves as an Associate Editor for the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS and *Complex and Intelligent Systems*.

Sam Kwong (Fellow, IEEE) received the B.Sc. degree in electrical engineering from the State University of New York at Buffalo, Buffalo, NY, USA, in 1983, the M.A.Sc. degree in electrical engineering from the University of Waterloo, Waterloo, ON, Canada, in 1985, and the Ph.D. degree in electrical engineering from the University of Hagen, Hagen, Germany, in 1996.

From 1985 to 1987, he was a Diagnostic Engineer with the Control Data Canada, Mississauga, ON, Canada, where he designed the diagnostic software to detect the manufacture faults of the VLSI chips in the Cyber 430 machine. He later joined the Bell Northern Research Canada, Ottawa, ON, Canada, as a Member of Scientific Staff. In 1990, he joined the City University of Hong Kong, Hong Kong, as a Lecturer with the Department of Electronic Engineering, where he is currently a Professor with the Department of Computer science. Since 2014, he has been the Vice President for IEEE Systems, Man and Cybernetics for conferences and meetings. His research areas are in pattern recognition, evolutionary computations, and video analytics.

Prof. Kwong was elevated to IEEE Fellow for his contributions on Optimization Techniques for Cybernetics and Video coding in 2014. He is also appointed as an IEEE Distinguished Lecturer for IEEE SMC society in March 2017.

Jie Zhang received the master's degree in computer science from the China University of Mining and Technology, Xuzhou, China, in 1999.

She is currently an Associate Professor with the Beijing University of Chemical Technology, Beijing, China. Her current research interests include formal verification, and PHM for power and embedded system design.

Hua Wang (Senior Member, IEEE) received the Ph.D. degree in computer science from the University of Southern Queensland, Toowoomba, QLD, Australia, in 2004.

He was a Professor with the University of Southern Queensland. He is currently a Full Time Professor with Victoria University, Melbourne, VIC, Australia. As a Chief Investigator, three Australian Research Council Discovery grants have been awarded since 2006, and 200 peer-reviewed scholar papers have been published. Six Ph.D. students have already graduated under his principal supervision. He has over ten years teaching and working experience in Applied Informatics with both enterprise and university. He has expertise in electronic commerce, business process modeling, and enterprise architecture.

Tianlong Gu received the M.Eng. degree in electronic machinery from Xidian University, Xi'an, China, in 1987, and the Ph.D. degree in industrial automation from Zhejiang University, Hangzhou, China, in 1996.

From 1998 to 2002, he was a Research Fellow with the School of Electrical and Computer Engineering, Curtin University of Technology, Perth, WA, Australia, and a Post-Doctoral Fellow with the School of Engineering, Murdoch University, Perth.

He is currently a Professor with the School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, China. His research interests include formal methods, data and knowledge engineering, software engineering, and information security protocol.

Huaqiang Yuan received the Ph.D. degree in computer software from Shanghai Jiao Tong University, Shanghai, China, in 1996.

He is currently a Professor with the School of Computer Science and Network Security, Dongguan University of Technology, Dongguan, China. His current research interests include computational intelligence and cyberspace security.

Jun Zhang (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the City University of Hong Kong, Hong Kong, in 2002.

He is currently a Visiting Scholar with Victoria University, Melbourne, VIC, Australia. His research interests include computational intelligence, cloud computing, data mining, and power electronic circuits. He has published over 200 technical papers in his research area.

Dr. Zhang was a recipient of the China National Funds for Distinguished Young Scientists from the National Natural Science Foundation of China in 2011 and the First-Grade Award in Natural Science Research from the Ministry of Education, China, in 2009. He is currently an Associate Editor of the IEEE TRANSACTION ON EVOLUTIONARY COMPUTATION and IEEE TRANSACTION ON CYBERNETICS.