

Estimation of Distribution Algorithm with Multivariate T -Copulas for Multi-Objective Optimization

Ying Gao, Lingxi Peng, Fufang Li, Miao Liu, Xiao Hu

Department of Computer Science and Technology, Guangzhou University, Guangzhou, China
Email: falcongao@sina.com.cn

Received September 10, 2012; revised October 10, 2012; accepted October 18, 2012

ABSTRACT

Estimation of distribution algorithms are a class of evolutionary optimization algorithms based on probability distribution model. In this article, a Pareto-based multi-objective estimation of distribution algorithm with multivariate T -copulas is proposed. The algorithm employs Pareto-based approach and multivariate T -copulas to construct probability distribution model. To estimate joint distribution of the selected solutions, the correlation matrix of T -copula is firstly estimated by estimating Kendall's τ and using the relationship of Kendall's τ and correlation matrix. After the correlation matrix is estimated, the degree of freedom of T -copula is estimated by using the maximum likelihood method. Afterwards, the Monte Carlo simulation is used to generate new individuals. An archive with maximum capacity is used to maintain the non-dominated solutions. The Pareto optimal solutions are selected from the archive on the basis of the diversity of the solutions, and the crowding-distance measure is used for the diversity measurement. The archive gets updated with the inclusion of the non-dominated solutions from the combined population and current archive, and the archive which exceeds the maximum capacity is cut using the diversity consideration. The proposed algorithm is applied to some well-known benchmark. The relative experimental results show that the algorithm has better performance and is effective.

Keywords: Estimation of Distribution Algorithm; Pareto-Based Approach; T -Copulas; Multi-Objective Optimization

1. Introduction

Optimization problems are widely encountered in various fields of science and technology. When an optimization problem involves only one objective function, the task of obtaining the optimal solution is called single objective optimization. When an optimization problem involves more than one objective function, the task of finding one or more optimum solutions is called multi-objective optimization. Many real-world optimization problems involve multiple objectives. The fundamental difference between single-objective and multi-objective optimization lies in the cardinality of the optimal set. In problems with two or more conflicting objectives, there is no single optimum solution. There exist a number of solutions which are optimal. Such solutions are called Pareto-optimal solutions. The solution to a multi-objective optimization problem consists of a solution set with multiple solutions that may produce tradeoffs between different objectives. These solutions are called non-dominated solutions and the corresponding solution set is called the Pareto-front [1]. None of the solutions in the Pareto-front is best with respect to all objectives. In addition, no solution in the Pareto-front is better than any other solution in

the front with respect to all objectives. Hence, without any additional problem-specific information about the priorities of various objectives, all the solutions in the Pareto-front are important. The main objective of multi-objective optimization is to find many such solutions which reflect the tradeoffs between the objectives. Multi-objective optimization algorithms, especially those based on evolutionary principles, have been widely used to solve problems with multiple objectives. In recent years, a considerable amount of interest has been shown in multi-objective evolutionary algorithm (MOEA) and a number of different MOEAs have been suggested, such as Strength Pareto Evolutionary Algorithm (SPEA2) [2], Non-dominated Sorting Genetic Algorithm (NSGA-II) [3], and Fast Pareto Genetic Algorithm (FastPGA) [4].

Estimation of distribution algorithms (EDAs) [5] are a class of evolutionary algorithms based on probability distribution model. They estimate a probability distribution from population of solutions, and sample it to generate the next population. It has been proven that EDA has some special characteristics of good convergence, high efficiency, concise concept and been successfully extended to multi-objective optimization problems. The performance of an EDA highly depends on how well it

estimates and samples the probability distribution. A wide variety of EDAs using different techniques to estimate and sample the probability distribution have been proposed and are the subject of active research. Many EDAs use probabilistic graphical modeling techniques [6, 7] for this purpose. However, EDA using probabilistic graphical modeling techniques generally spend too much time on the learning about the probability distribution of the promising individuals. Copula [8,9] is a recently developed mathematical theory and a power tool for multivariate probability analysis. According to copula theory, a joint probability distribution can be decomposed into n marginal probability distributions and a copula function. So, the joint probability distribution of multivariate can be constructed utilizing a copula function and the marginal probability distributions of every variable. Copulas have been applied to constitute the probabilistic model in the conventional EDAs [10,11]. It is simpler and easier to model the probability distribution compared with other methods in EDAs.

In this paper, EDA is extended to multi-objective optimization problems by using a Pareto-based approach and T -copulas. The extended multi-objective EDA employs multivariate T -copulas to construct probability distribution model. T -copula parameters are firstly estimated, thus, joint distribution is estimated by estimating Kendall's τ and using the relationship of Kendall's τ and correlation matrix. After the correlation matrix is estimated, the degree of freedom of T -copula is estimated by using the maximum likelihood method. Afterwards, the Monte Carlo simulation is used to generate new individuals. An archive with maximum capacity is used to maintain the non-dominated solutions. The Pareto optimal solutions are selected from the archive on the basis of the diversity of the solutions, and the crowding-distance measure is used for the diversity measurement. The archive gets updated with the inclusion of the non-dominated solutions from the combined population and current archive, and the archive which exceeds the maximum capacity is cut using the diversity consideration. The proposed algorithm is applied to some well-known benchmark. The relative experimental results show that the algorithm has better performance and is effective.

2. Multi-objective Optimization Problems

The general multi-objective optimization problem can be defined as follows:

$$\begin{aligned} \min \mathbf{F}(\mathbf{x}) &= \{f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_k(\mathbf{x})\} \\ \text{s.t. } g_i(\mathbf{x}) &\leq 0, \quad i = 1, 2, \dots, m \end{aligned} \quad (1)$$

$$h_i(\mathbf{x}) = 0, \quad i = 1, 2, \dots, p \quad (2)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbf{X}$ is an n -dimensional deci-

sion variable vector and \mathbf{X} is the decision variable space. The constraints given by (1) and (2) define the feasible region Ω and any point in Ω defines a feasible solution. $M = \{\mathbf{y} | \mathbf{y} = \mathbf{F}(\mathbf{x}), \mathbf{x} \in \Omega\}$ is referred to as objective space. The k components of the vector $\mathbf{F}(\mathbf{x})$ are the criteria to be considered. The constraints $g_i(\mathbf{x})$ and $h_i(\mathbf{x})$ represent the restrictions imposed on the decision variables.

When there are several objective functions, the concept of optimum changes, because in multi-objective optimization problems the purpose is to find "trade-off solutions rather than a single solution. The concept of optimum commonly adopted in multi-objective optimization is Pareto optimality. Pareto optimality is defined as:

A point $\mathbf{x}^* \in \Omega$ is Pareto optimal if $\forall \mathbf{x} \in \Omega$ and $I = \{1, 2, \dots, k\}$ either: $\forall i \in I \quad f_i(\mathbf{x}^*) \leq f_i(\mathbf{x})$ and, there is at least one $i \in I$ such that $f_i(\mathbf{x}^*) < f_i(\mathbf{x})$

This definition says that \mathbf{x}^* is Pareto optimal if there exists no feasible vector \mathbf{x} which would decrease some criteria without causing a simultaneous increase in at least one other criterion. Other important definitions associated with Pareto optimality are Pareto dominance.

A vector $\mathbf{x} = (x_1, x_2, \dots, x_n)$ is said to dominate $\mathbf{y} = (y_1, y_2, \dots, y_n)$, denoted by $\mathbf{x} < \mathbf{y}$, if and only if \mathbf{x} is partially less than \mathbf{y} , i.e., $\forall i \in \{1, 2, \dots, k\} \quad x_i \leq y_i$ and, at least for one i , $x_i < y_i$.

For a given multi-objective problem $\mathbf{F}(\mathbf{x})$, the Pareto optimal set P^* is defined as:

$$P^* = \{\mathbf{x} \in \Omega | \neg \exists \mathbf{x}' \in \Omega \quad \mathbf{F}(\mathbf{x}') < \mathbf{F}(\mathbf{x})\}$$

For a given multi-objective problem $\mathbf{F}(\mathbf{x})$ and Pareto optimal set P^* , the Pareto front PF^* is defined as:

$$PF^* = \{\mathbf{F}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_k(\mathbf{x})) | \mathbf{x} \in P^*\}$$

The set of all Pareto optimal solutions in the feasible region Ω is called Pareto optimal set and the corresponding set of objective vector is called Pareto optimal front. The illustrative example of a multi-objective minimization problem with two objectives, f_1 and f_2 , that are plotted in the objective space M mapped from the feasible region Ω is shown in **Figure 1**. The bold curve in the feasible region Ω indicates the Pareto optimal set P^* . The bold curve in the objective space M indicates the Pareto front PF^* .

3. Multivariate T -Copulas

Estimation of distribution algorithms differ from traditional evolutionary algorithms. Instead of applying genetic operators like mutation and crossover to the parents, estimation of distribution algorithms estimate a probability distribution over the search space based on how the parent individuals are distributed in the search space, and then sample the offspring individuals from this distribu-

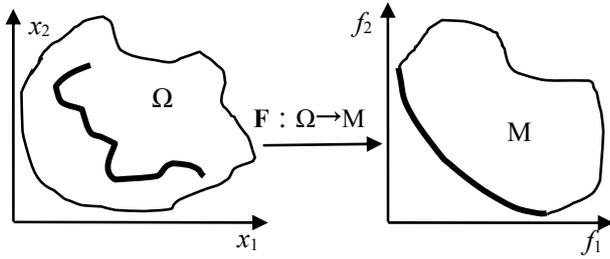


Figure 1. The illustrative example.

tion. The major issues in estimation of distribution algorithms are how to build a probability distribution model and how to sample the new individuals according to the probability distribution model.

The theory of copulas is known to provide a useful tool for modelling dependence in many applications. Copulas have attracted significant attention in the recent literature for modeling probability distribution of multivariate observations. An important feature of copulas is that they enable us to specify the univariate marginal distributions and their joint behavior separately. Archimedean Copula and Gaussian Copula have been applied to constitute the probabilistic model in EDAs [10,11].

The copula theory is briefly introduced in the following, the details can be found in [8,9].

An n -dimensional copula is a multivariate C.D.F., C , with uniformly distributed margins on $[0, 1]$ ($U(0,1)$) and the following properties:

- 1) $C: [0, 1]^n \rightarrow [0, 1]$;
- 2) C is grounded and n -increasing;
- 3) C has margins C_i which satisfy

$$C_i(u) = C(1, \dots, 1, u, 1, \dots, 1) = u \text{ for all } u \in [0, 1].$$

It is obvious, from the above definition, that if F_1, \dots, F_n are univariate distribution functions,

$C(F_1(x_1), \dots, F_n(x_n))$ is a multivariate C.D.F. with margins F_1, \dots, F_n , since $U_i = F_i(X_i)$, $i = 1, \dots, n$, is a uniform random variable. Copula functions are a useful tool to construct and simulate multivariate distributions.

The following theorem is known as **Sklar's Theorem**. It is the most important theorem regarding to copula functions since it is used in many practical applications.

Theorem: Let F be an n -dimensional C.D.F. with continuous margins F_1, \dots, F_n . Then F has the following unique copula representation (canonical decomposition):

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)) \quad (3)$$

The theorem of Sklar is very important, because it provides a way to analyse the dependence structure of multivariate distributions without studying marginal distributions. From Sklar's theorem we see that, for continuous multivariate distribution functions, the univariate margins and the multivariate dependence structure can be

separated. The dependence structure can be represented by an adequate copula function. Moreover, the following corollary is attained from (3).

Corollary: Let F be an n -dimensional C.D.F. with continuous margins F_1, \dots, F_n and copula C (satisfying (3)). Then, for any $u = (u_1, \dots, u_n)$ in $[0, 1]^n$:

$$C(u_1, \dots, u_n) = F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)) \quad (4)$$

where F_i^{-1} is the generalized inverse of F_i .

The T -copula can be thought of as representing the dependence structure implicit in a multivariate T -distribution. For a symmetric and positive definite matrix $R = (r_{i,j}), i, j = 1, 2, \dots, n$, with unit diagonal entries, let $T_{R,v}$ denote the standardized multivariate Student's T -distribution with correlation matrix R and $v \geq 1$ degrees of freedom:

$$T_{R,v}(y_1, \dots, y_n) = \int_{-\infty}^{y_1} \dots \int_{-\infty}^{y_n} \frac{\Gamma\left(\frac{v+n}{2}\right) |R|^{-1/2}}{\Gamma\left(\frac{v}{2}\right) (v\pi)^{n/2}} \times \left(1 + \frac{1}{v} \mathbf{x}^T R^{-1} \mathbf{x}\right)^{-(v+n)/2} dx_1 \dots dx_n \quad (5)$$

The multivariate T -copula (MTC) is defined as:

$$C(u_1, u_2, \dots, u_n; R, v) = T_{R,v}(T_v^{-1}(u_1), \dots, T_v^{-1}(u_n))$$

where T_v^{-1} is the inverse of the univariate Student's t cumulative distribution function with v degrees of freedom.

The corresponding density is

$$c(u_1, u_2, \dots, u_n; R, v) = \frac{1}{|R|^{1/2}} \frac{\Gamma\left(\frac{v+n}{2}\right)}{\Gamma\left(\frac{v}{2}\right)} \left(\frac{\Gamma\left(\frac{v}{2}\right)}{\Gamma\left(\frac{v+1}{2}\right)}\right)^n \times \frac{\left(1 + \frac{1}{v} \boldsymbol{\omega}^T R^{-1} \boldsymbol{\omega}\right)^{-(v+n)/2}}{\prod_{j=1}^n (1 + \omega_j^2/v)^{-(v+1)/2}} \quad (6)$$

with $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_n)^T$; $\omega_j = T_v^{-1}(u_j), j = 1, \dots, n$.

A number of recent papers have shown that the empirical fit of the T -copula is generally superior to that of the so-called Gaussian copula, the dependence structure of the multivariate normal distribution. One reason for this is the ability of the T -copula to capture better the phenomenon of dependent extreme values. In the paper, we use T -copulas to model probability distribution in multi-objective estimation of distribution algorithm.

4. Multi-Objective EDA with T -Copulas

EDAs are undoubtedly a powerful search engine for solv-

ing single objective optimization problems. However, the original scheme has to be modified for solving multi-objective optimization problems. As we saw in Section II, the solution set of a problem with multiple objectives does not consist of a single solution. Instead, in multi-objective optimization, we aim to find Pareto optimal set. Besides using multivariate T -copulas for constructing probability distribution model in the multi-objective EDA, we have used an archive with maximum capacity to maintain the non-dominated solutions and the crowding-distance measure for the diversity measurement. The Pareto optimal solutions are selected from the archive on the basis of the diversity of the solutions. The archive gets updated with the inclusion of the non-dominated solutions from the combined population and current archive, and the archive which exceeds the maximum capacity is cut using the diversity consideration. The algorithm steps of multi-objective EDA with T -copulas for solving multi-objective optimization problems are as follows:

```

t=0, initialize a population
S(0) ← Initialization ();
Evaluate(S(0))
A(0) ← Non_Dominated(S(0))
WHILE t < T DO
  Estimate_Marginals_Distribution();
  Estimate_Correlation_Matrix();
  Estimate_Degrees_Freedom();
  S'(t) ← Sample();
  S(t+1) ← Selection(S(t) ∪ S'(t))
  Evaluate(S(t+1))
  A(t+1) ← Non_Dominated(S(t+1) ∪ A(t))
  IF |A(t+1)| > M THEN Cut_Archive(A(t+1))
  t ← t+1
END WHILE
Output obtained Pareto optimal front

```

The function initialization() is used to initialize a population $S(0)$ with size N . The initial population $S(0)$ is chosen randomly and should uniformly cover the entire solution space based on the consideration of the requirement of population diversity. The function Evaluate() is used to evaluate each individual in population. The archive $A(0)$ has been initialized to contain the non-dominated solutions from $S(0)$. The function Non_Domina() returns the non-dominated solutions from a population.

The function Estimate_Marginals_Distribution() is used to estimate the marginals distribution from population. The function Estimate_Correlation_Matrix() is used to estimate the correlation matrix \mathbf{R} in T -copulas.

A simple method [9] based on Kendall's τ is applied for estimating the correlation matrix \mathbf{R} . The method consists of constructing an empirical estimate of Kendall's

τ for each bivariate margin of the copula and then using relationship (7)

$$\tau(\mathbf{X}_k, \mathbf{X}_m) = \frac{2}{\pi} \arcsin(r_{k,m}) \quad (7)$$

To infer an estimate of the relevant element of \mathbf{R} . More specifically we estimate $\tau(\mathbf{X}_k, \mathbf{X}_m)$ by calculating the standard sample c coefficient

$$\hat{\tau}(\mathbf{X}_k, \mathbf{X}_m) = \frac{2}{n(n-1)} \sum_{1 \leq i < j \leq n} \text{sign}[(X_{i,k} - X_{j,k})(X_{i,m} - X_{j,m})] \quad (8)$$

From the original data vectors $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$, and write the j th component of the i th vector as $X_{i,j}$; this yields an unbiased and consistent. An estimator of $r_{k,m}$

is then given by $\hat{r}_{k,m} = \sin\left(\frac{\pi}{2} \hat{\tau}(\mathbf{X}_k, \mathbf{X}_m)\right)$. In order to

obtain an estimator of the entire matrix \mathbf{R} we can collect all pairwise estimates $\hat{r}_{k,m}$ and then construct the estimator

$$\hat{\mathbf{R}} = \left(\sin\left(\frac{\pi}{2} \hat{\tau}(\mathbf{X}_k, \mathbf{X}_m)\right) \right), k=1, \dots, N, m=1, \dots, N \quad (9)$$

Then, The function Estimate_Degrees_Freedom() is used to estimate the degrees of freedom ν by maximum likelihood method.

The function Sample() is used to generate offspring population $S'(t)$ by sampling a T -copula function $T_{\mathbf{R},\nu}$ with correlation matrix \mathbf{R} and ν degrees of freedom can be generated as follows [9].

- 1) Find the Cholesky decomposition of \mathbf{R} , so that $\mathbf{A}\mathbf{A}^T = \mathbf{R}$, with \mathbf{A} lower triangular;
- 2) Generate a sample of n independent random variables $\mathbf{Z} = (z_1, z_2, \dots, z_n)^T$ from $N(0, 1)$;
- 3) Generate an independent random variables s with Z from χ^2_ν distribution;

$$4) \text{ Set } \mathbf{Y} = \mathbf{A}\mathbf{Z}, \mathbf{W} = \frac{\sqrt{\nu}}{\sqrt{s}} \mathbf{Y} \text{ with}$$

$$\mathbf{W} = (w_1, w_2, \dots, w_n)^T \text{ and } \mathbf{Y} = (y_1, y_2, \dots, y_n)^T;$$

$$5) \text{ Set } u_j = T_\nu(w_j), j=1, \dots, n;$$

6) Return

$$\mathbf{X} = (x_1, x_2, \dots, x_n) = (F_1^{-1}(u_1), F_2^{-1}(u_2), \dots, F_n^{-1}(u_n)).$$

The function Selection() is used to evaluate each individual in the current population $S(t)$ and the offspring population $S'(t)$ and carry out the non-dominated sorting and ranking selection, developed by Deb *et al.* in [3] to select non-dominated individuals to constitute the next generation population $S(t+1)$.

The proposed algorithm maintains an archive $A(t)$ with maximum capacity M . At each iteration, the archive gets updated with the inclusion of the non-dominated

solutions from the next generation population $S(t+1)$ and the archive $A(t)$. If the size of the archive exceeds the maximum capacity M , it is cut using the diversity consideration. The function `Cut_Archive()` is used to cut the archive.

5. Simulation Results

In order to test efficiency of the Pareto-based multi-objective estimation of distribution algorithm with T -copulas, the performance of the proposed algorithm is compared with that of some multi-objective optimization algorithms. The following some well-known benchmark functions [12] have been used to test.

1) **ZDT1:**

$$f_1(\mathbf{x}) = x_1$$

$$f_2(\mathbf{x}) = g(\mathbf{x})\left(1 - \sqrt{x_1/g(\mathbf{x})}\right) \quad x_i \in [0,1]$$

$$g(\mathbf{x}) = 1 + 9\left(\sum_{i=2}^D x_i\right)/(D-1)$$

2) **ZDT2:**

$$f_1(\mathbf{x}) = x_1$$

$$f_2(\mathbf{x}) = g(\mathbf{x})\left(1 - (x_1/g(\mathbf{x}))^2\right) \quad x_i \in [0,1]$$

$$g(\mathbf{x}) = 1 + 9\left(\sum_{i=2}^D x_i\right)/(D-1)$$

3) **ZDT3:**

$$f_1(\mathbf{x}) = x_1$$

$$f_2(\mathbf{x}) = g(\mathbf{x})\left(1 - \sqrt{\frac{x_1}{g(\mathbf{x})}} - \frac{x_1}{g(\mathbf{x})}\sin(10\pi x_1)\right)$$

$$x_i \in [0,1]$$

$$g(\mathbf{x}) = 1 + 9\left(\sum_{i=2}^D x_i\right)/(D-1)$$

4) **ZDT4:**

$$f_1(\mathbf{x}) = x_1$$

$$f_2(\mathbf{x}) = g(\mathbf{x})\left(1 - \sqrt{x_1/g(\mathbf{x})}\right)$$

$$x_1 \in [0,1], x_i \in [-5,5],$$

$$i = 2, \dots, D$$

$$g(\mathbf{x}) = 1 + 10(D-1) + \sum_{i=2}^D (x_i^2 - 10\cos(4\pi x_i))$$

5) **ZDT6:**

$$f_1(\mathbf{x}) = 1 - \exp(-4x_1)\sin^6(6\pi x_1)$$

$$f_2(\mathbf{x}) = g(\mathbf{x})\left(1 - (f_1(\mathbf{x})/g(\mathbf{x}))^2\right)$$

$$x_i \in [0,1]$$

$$g(\mathbf{x}) = 1 + 9\left(\sum_{i=2}^D x_i/(D-1)\right)^{0.25}$$

The two common metrics used to compare are convergence metric γ and divergence metric Δ . For these metrics we need to know the true Pareto front for a problem. In our experiments we use 1000 uniformly spaced Pareto optimal solutions as the approximation of the true Pareto front. A brief introduction of these metrics is given here:

Convergence metric γ was proposed by Deb *et al.* [3], measures the distance between the obtained non-dominated front NF and optimal Pareto front PF . Mathematically, it may be defined as:

$$\gamma = \frac{\sqrt{\sum_{i=1}^N d_i}}{N} \quad (10)$$

where N is the number of non-dominated solutions found by the algorithm being analyzed and d_i is the minimum Euclidean distance (measured in the objective space) between the i th solution of NF and the solutions in PF .

Diversity metric Δ was proposed by Deb *et al.* [3]. It measures the extent of spread achieved among the obtained solutions and is defined as:

$$\Delta = \frac{d_f + d_i + \sum_{i=1}^{s-1} |d_i - \bar{d}|}{d_f + d_i + (s-1)\bar{d}} \quad \text{where } \bar{d} = \frac{\sum_{i=1}^{s-1} d_i}{s-1} \quad (11)$$

where s is the number of members in the set of non-dominated solution found so far, and the parameter d_f and d_i are the Euclidean distances between the extreme solutions and the boundary solutions of the obtained non-dominated set. The parameter \bar{d} is the average of all distances d_i ($i=1, \dots, s-1$), assuming that there are s solutions on the best non-dominated front. The smaller the value of these metrics is, the better the performance of the algorithm is. The initial population was generated from a uniform distribution in the ranges specified below. Population size $N = 80$. All experiments were repeated for 30 runs. The maximum number of iterations is set to 1000 in each running.

Table 1 listed the mean and standard deviation(sd) values of convergence metric γ and diversity metric Δ obtained using SPEA2 [2], NSGA-II [3], FastPGA [4] and the proposed algorithm(EDATCMO) on the multi-objective benchmark functions ZDT1, ZDT2, ZDT3, ZDT4 and ZDT6. It can be known from **Table 1**, the mean values of convergence metric γ obtained by using EDATCMO is smaller than those obtained by using SPEA2, NSGA-II, FastPGA. And the mean values of diversity metric Δ obtained by using EDATCMO is smaller

Table 1. Mean and sd of the convergence and diversity metrics for Benchmark functions.

Problem	Algorithm	Convergence γ (mean \pm sd)	Diversity Δ (mean \pm sd)
ZDT1	SPEA2	0.09462 \pm 0.04511	0.42209 \pm 0.01012
	NSGA-II	0.10872 \pm 0.00362	0.50827 \pm 0.02446
	FastPGA	0.09154 \pm 0.00621	0.70009 \pm 0.01174
ZDT2	EDATCMO	0.02363 \pm 0.00146	0.21738 \pm 0.00368
	SPEA2	0.08073 \pm 0.06101	0.50013 \pm 0.01612
	NSGA-II	0.09023 \pm 0.00401	0.30163 \pm 0.01503
ZDT3	FastPGA	0.02067 \pm 0.00702	0.60034 \pm 0.02984
	EDATCMO	0.00821 \pm 0.00472	0.20022 \pm 0.00863
	SPEA2	0.13996 \pm 0.07603	0.74932 \pm 0.05006
ZDT4	NSGA-II	0.04208 \pm 0.09104	0.85061 \pm 0.09025
	FastPGA	0.07024 \pm 0.06541	0.86027 \pm 0.06453
	EDATCMO	0.02005 \pm 0.01843	0.50064 \pm 0.06132
ZDT6	SPEA2	0.62768 \pm 0.10676	0.51682 \pm 0.07319
	NSGA-II	1.13458 \pm 0.90054	0.85454 \pm 0.09003
	FastPGA	2.97402 \pm 1.94585	0.96856 \pm 0.07032
	EDATCMO	0.50547 \pm 0.08019	0.41621 \pm 0.02971
	SPEA2	0.91036 \pm 0.09105	0.74532 \pm 0.04093
	NSGA-II	0.49311 \pm 0.00874	0.63847 \pm 0.08006
	FastPGA	0.72347 \pm 0.01106	0.84442 \pm 0.04034
	EDATCMO	0.10048 \pm 0.01002	0.31618 \pm 0.01782

than those obtained by using SPEA2, NSGA-II, FastPGA. It indicates that EDATCMO outperforms SPEA2, NSGA-II and FastPGA algorithms on the benchmark functions in both aspects of convergence and distribution of solutions.

6. Conclusion

We proposed a multi-objective estimation of distribution algorithm using T -copulas and Pareto-based approach. The proposed algorithm employs the multivariate T -copulas to construct probability distribution model, and the new individuals are generated according to the probability distribution model. An archive is used to maintain the non-dominated solutions. The Pareto optimal solutions are selected from the archive. The archive gets updated with the inclusion of the non-dominated solutions from the combined population and current archive. The algorithm is applied to some well-known benchmarks. The results show that the algorithm has better performance than SPEA2 [2], NSGA-II [3], FastPGA [4] on the multi-objective benchmark functions ZDT1, ZDT2, ZDT3, ZDT4 and ZDT6.

7. Acknowledgements

This work is supported by the Scientific and Technological Innovation Projects of Department of Education of Guangdong Province, P.R.C. Grant No. 2012KJCX0082 and Science and Technology Projects of Guangdong Province, P.R.C. under Grant No. 2011B090400623, Guangzhou Science and Technology Projects under Grant No. 12C42011563, 11A11020499.

REFERENCES

- [1] C. A. Coello, "A Comprehensive Survey of Evolutionary-Based Multi-objective Optimization Techniques," *International Journal of Knowledge and Information Systems*, Vol. 1, No. 3, 1999, pp. 269-308.
- [2] E. Zitzler, M. Laumanns and L. Thiele, "SPEA2: Improving the Strength Pareto Evolutionary Algorithm for Multi-Objective Optimization," *Proceedings of the EUROGEN Conference*, Lake Como, 2001, pp. 95-100.
- [3] K. Deb, S. Agrawal, A. Pratap and T. Meyarivan, "A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, Vol. 6, No. 2, 2002, pp. 182-197. [doi:10.1109/4235.996017](https://doi.org/10.1109/4235.996017)
- [4] H. Eskandari, C. D. Geiger and G. B. Lamont, "FastPGA: A Dynamic Population Sizing Approach for Solving Expensive Multi-Objective Optimization Problems," In: *Evolutionary Multi-objective Optimization Conference EMO*, Springer-Verlag, Berlin, Heidelberg, 2007, pp. 141-155.
- [5] P. Larranaga and J. A. Lozano, "Estimation of Distribution Algorithms: A New Tool for Evolutionary Computation," Kluwer Academic Publishers, Dordrecht, 2002. [doi:10.1007/978-1-4615-1539-5](https://doi.org/10.1007/978-1-4615-1539-5)
- [6] S. Shakya, "DEUM: A Framework for an Estimation of Distribution Algorithm Based on Markov Random Fields," Ph.D. thesis, The Robert Gordon University, Aberdeen, 2006.
- [7] S. Shakya and J. McCall, "Optimisation by Estimation of Distribution with DEUM Framework Based on Markov Random Fields," *International Journal of Automation and Computing*, Vol. 4, No. 3, 2007, pp. 262-272. [doi:10.1007/s11633-007-0262-6](https://doi.org/10.1007/s11633-007-0262-6)
- [8] R. B. Nelsen, "An Introduction to Copula," Springer-Verlag, New York, 1998.
- [9] U. Cherubini, E. Luciano and W. Vecchiato, "Copula Methods in Finance," John Wiley & Sons Ltd., Chichester, 2004.
- [10] Y. Gao, "Multivariate Estimation of Distribution Algorithm with Laplace Transform Archimedean Copula," *IEEE International Conference on Information Engineering and Computer Science*, Vol. 1, 2009, pp. 273-277.
- [11] Y. Gao, X. Hu, H. L. Liu, F. F. Li and L. X. Peng, "Opposition-Based Learning Estimation of Distribution Algorithm with Gaussian Copulas and Its Application to Placement of RFID Readers," *International Conference on Artificial Intelligence and Computational Intelligence*, Taiyuan, 24-25 September 2011, pp. 219-227.

- [12] E. Zitzler, K. Deb and L. Thiele, "Comparison of Multi-Objective Evolutionary Algorithms: Empirical Results,"

Evolutionary Computation, Vol. 8, No. 2, 2000, pp. 173-195. [doi:10.1162/106365600568202](https://doi.org/10.1162/106365600568202)