Knot Estimation of the B-Spline Curve with Strength Pareto Evolutionary Algorithm 2 (SPEA2)

SABAN GÜLCÜa and ERKAN ÜLKERb
Computer Engineering Department
Konya Universitya and Selçuk Universityb
Alaeddin Keykubat Campus, Selçuklu, Konya
TURKEY
saban.gulcu@karatay.edu.tr, eulker@selcuk.edu.tr

Abstract: - Real-world problems are often multi-objectives. For solve these problems, multiple objectives conflicting with each other must simultaneously be optimize. Multi-objective optimization is very different from single-objective optimization. The objective for single-objective is to obtain a single design or the best decide. In the objectives conflicting with each other in a multi-objective optimization, the consensus on the set of obtained optimal solutions are provided. This set is often called to as a set of Pareto-optimal solutions. If a set of irregular points assumed to lie on or passes close to an unknown curve is given, B-spline curve estimation is a multi-objective Combinatorial optimization problem for the aim of finding parameters (control points, knot vector, etc) of the B-spline curve to estimates the corresponding curve in the order. Since 1985 for solving multi-objective optimization, several approaches using evolutionary algorithms are proposed. In this article, the B-spline curve knots has been estimated by using Strength Pareto Evolutionary Algorithm 2. The success and the performance of the method are shown by the tests.

Key-Words: - Knot optimization, B-Spline Curve, Approximation, Strength Pareto Evolutionary Algorithm 2

1 Introduction
Combinatorial optimization problem is defined by a set of variables, variable domains, variables, constraints, and objective function to be minimized. Referred to as the search space where space of possible solutions. While the solution area of each variable represent as domain, each solution named as candidate solution. for the solve a combinatorial optimization problem, the global optimal solution provide the minimum objective function should be searched. The term of the metaheuristic proposed by the Greek and was a combination of two word [1]. The heuristic word derived from search verb, the word meta means "beyond, in an upper level”. These methods are also referred to as the modern heuristic [2]. The most popular metaheuristic algorithms are follows; Ant Colony Optimization (ACO), Artificial Bee Colony (ABC), Evolutionary Computation (EC), including Genetic Algorithms (GA), Iterated Local Search (ILS), Simulated Annealing (SA), and Tabu Search (TS).

The concept of Pareto optimal in the early 1900s, originally founded by an Italian economist and sociologist Vilfredo Pareto who comes to the economic theory of the new prosperity. While Pareto search for the possible equilibrium states of a nation's economy the name itself has found that the concept of Pareto optimal. To reach Pareto optimal situation where people with useful goods and services requirements, would have been satisfied with the best possible way. With this, the foundations of the theory of modern welfare founded. Pareto-optimal concept, to adaptation evolutionary algorithms, multi-objective evolutionary optimization algorithms is developed. Accordingly, the solution in the population according to the aims, the best and the worst, and may be identical to the other solutions. The best solution for the worst any one of the objectives are not, at least meaning solution is better than others for a particular objective. The optimal solution, quenched by any other solution in the search space is the solution. While such a optimal solution is named as Pareto-optimal solution, in such a whole set of solutions for such Pareto-optimal trade-offs called Optimal cluster [3-5]. The concept of Pareto optimality or Pareto-based approaches, tried to be achieved should not be inhibited by using the classification based on the rule. Algorithms using this approach, rating method inspired by Pareto in 1989 revealed of Golberg[6]. Goldberg suggests, the current population, which is inhibited by the appropriate solutions, using a selection mechanism.
that tells the moving population toward the surface of Pareto. In addition, as the diversity of compliance, the use of techniques the sharing and niching recommends. In recent years, a wide variety of Pareto-based MOEA developed. Major multi-objective algorithms can be given follows: genetic algorithms, Vector Evaluated Genetic Algorithm, Multiobjective Genetic Algorithm, Niched Pareto Genetic Algorithm, Nondominated Sorting Genetic Algorithm, Nondominated Sorting Genetic Algorithm-II, Strength Pareto Evolutionary Algorithm, Strength Pareto Evolutionary Algorithm 2, and Pareto Envelope-based Selection Algorithm.

In this area, VEGA of Schaffer [7] is pioneer working and to produce good results for simple problems, which is primitive population-based optimization algorithm. Similarly, Muratta's MOGA [8] is the population-based multi-objective evolutionary algorithm too and shows a good distribution than VEGA. But the reality is that a distribution of optimal solutions, remote and poor profit results for the other algorithms produced by perfect match, the results are quite inefficient. Goldberg's Pareto surface is degread and in the same year the proposal put forward by addressing one of the algorithms Horn and Nafpliotis's NPGA [9]. NPGA uses Pareto superiority tournament in selection process. In this method, a certain number of individuals and a comparison of population research using a set of randomly selected as the winner of the tournament is done and the individual selected. The NSGA algorithm developed by Srinivas and Deb [10] and which was based on Pareto superiority is one of the successful algorithms. Goldberg's suggestion in the NSGA, sorting the population according to Pareto superiority fake fitness values assigned to them and it presents various uses to provide niching technique performed more efficiently way. However, the high computational complexity of this algorithm, the lack of exclusivity and such as there is weak point value of the parameter $\sigma$ share configuration uncertainty. Considering these deficiencies in the algorithm improved by Deb et.al. and have been suggested as NSGAII [11], SPEA, both in terms of the diversity of the Pareto surface and the convergence of solutions as well as more successful from VEGA, MOGA, NPGA, and NSGA. However, have several drawbacks like assignment eligibility, density estimation, and reducing the size of archive. SPEA2 has been developed to eliminate them.

B-spline curves in the industry, computer-based design, computer based manufacturing, computer graphics is a commonly used mathematical interpolation and forecasting and is a type of parametric curve. By its very nature the ability to represent the curve of the best spots and, correspondingly, the knot is marked as a vector of knots produced. After estimate knot vector, the control points calculated. The B-spline curve are trying to reproduce a set of distributed points via Knot vector, control points, and blending functions. The predicted curve is expected to be the minimum amount of error between the actual curve. Therefore, the objective function minimizes the error function and the objective function variables are B-spline coefficients and the internal knots. B-spline coefficients are linear parameters. However, the inner knots are the nonlinear parameters, because the predicted function of B-spline curve is a nonlinear function of the knots. This minimization problem is known as a Combinatorial multimodal optimization problem. In this paper, the B-spline curve is estimated by the Strength Pareto Evolutionary Algorithm 2 algorithm, which is one of the different variants of genetic algorithm metaheuristic. Organization of this article is as follows. B-spline curve estimation problem are detailed in subsequent chapter. After giving details Strength Pareto Evolutionary Algorithm 2 (SPEA2), the process of the B-spline curve knot estimation by using SPEA2 is given. After viewing the results and success of the performance of proposed algorithm conclusion section ended with the article.

2 B-Spline Curve Estimation Problem

Bezier curves are parametric curves, one of the first species to overcome the problem despite numerous detailed modeling has some limitations and disadvantages. To overcome the disadvantages of this, type of B-spline curves have a local domain uses blending functions. In other words, the blending functions of the B-spline curve is equal to zero outside parts of them. Therefore, the shape of the curve is just adjacent to it a few control points. B-spline curves by Bezier curves is another important advantage is the B-spline curve is independently of the degree of control points. Despite these advantages Bezier curves, programming of B-spline curves are more difficult and complicated. Mathematical representation of B-spline curves look like Bezier curves and as follows:

$$P(u) = \sum_{i=0}^{n} p_i N_{i,d}(u) \quad (1)$$

Despite this similarity, for the blending function is different, between Bezier and B-spline curves leads to the following main differences:
first, \( u \) cannot take the parameter values in [0,1]. One other difference is that \( N_{i,d} \) as previously mentioned, the degree of the blending functions is not dependent on the number of control points.

Cox-deBoor recursive functions forms the basis of B-spline curves and B-spline blending functions calculated using the Cox-deBoor functions:

\[
N_{i,1}(u) = \begin{cases} 1, & t_i \leq u \leq t_{i+1} \\ 0, & \text{otherwise} \end{cases} 
\]

\[
N_{i,k}(u) = \frac{u-t_i}{t_{i+k-1} - t_i} N_{i,k-1}(u) + \frac{t_{i+k} - u}{t_{i+k} - t_{i+1}} N_{i+1,k-1}(u) 
\]

In the B-spline curves, there are three main features that define the curve; control points \( p_n \), the degree of curve \( (d) \) and the knot vector for \( u \) parameters.

General methods of reconstruction for B-spline curves can be edited as follows:

1) Order of Curve \((k=d+1)\), the number of control points \((n)\) and the knot values \( t_i \) is fixed.

2) Each point dispersed in a \( F_j \), is assigned to the parameter value \( C(t_i) \).

3) \( C(t_j) = F_j \) system is solved or \( \sum_j |c_j - f_j| \) minimized.

The crux of this strategy in terms of B-spline curves, often referred to as the second step in the parameterization of the data. For detailed information about the parametric curve fitting problem, readers may referred to Yang et.al. [12].

Proposed on the basis of the above strategy, the method followed here to study, detailed in the B-spline curve estimation. In B-spline curve and surface fitting, distributed data problem turn to a discrete combinatorial optimization problem by using the conversion process from mathematical foundation in benefit Yoshimoto et.al.[13], Safraz and Raza [14]’s studies. SPEA2 then converted by means of a new method is proposed to solve the problem. To measure the performance and quality of the method with the traditional genetic algorithm, a study has been conducted and the results were compared.

Invented data, interval \([a, b]\) in the x-axis is given in the range of indoor and have equation (2). In this equation, \( f(x) \) function (unknown) is the function on the basis of the data and \( \Delta j \)'s remember that a measurement error. \( t_i \) \((i=1-m, 2-m, \ldots, n+m)\) are knots of the curve. Where \( n \) is the letter \((a,b)\), which is located in the range of \( t_i \) \((i=1, 2, \ldots, n)\) . The number of knots, \( m \) is a term \( N_{m} \) B-spline's scheme (degree +1). \([a, b]\), closed at the end of the range of values of \( a \) and \( b \), the knots are synchronized.

\[
F_j = f(x_j) + \Delta_j, \quad (j = 1, 2, \ldots, N) \tag{3}
\]

\[
\begin{align*}
\{a &= t_{1-m} = \ldots t_0, \\
b &= t_{n+1} = \ldots t_{n+m} \}
\end{align*} \tag{4}
\]

The setting of \( f(x) \) for a model function is given as follows:

\[
S(x) = \sum_{j=1}^{n+m} c_j N_{m,j}(x) \tag{5}
\]

Where, the expression of \( c_j \) is B-spline coefficient. B-splines in Equation (3) to be calculated with equation (2). Equation (2), the knots at the same time not only the denominator of the fractions in the share. Therefore, equation (5) with a given spline, a nonlinear function of the knots. Equation (5), by means of the method of least squares and is adapted to the data in equation (3). Then, the remaining \( Q_1 \) is the sum of the squares of the equation (6) and is calculated. Here, the weight data \( W_i \) and \( N > n + m \). All of the weight values were \( l \). In this study, Centripetal method is used for determining a knot vector.

\[
Q_1 = \sum_{j=0}^{N} W_j \left[ S(x_j) - F_j \right]^2 
\]

\[
\text{As explained above, the minimizing objective function is the equation (6), and S(x) function is a nonlinear function of the knots. This minimization problem is known as a Combinatorial multimodal optimization problem.} 
\]

3 Strength Pareto Evolutionary Algorithm

Strength Pareto Evolutionary Algorithm proposed by Eckart Zitztler in 1999 (SPEA) [15], Pareto-optimal set by methods known to converge your old ones will alter spreading unite. SPEA’s main features are as follows: i) all individuals that have been covered so far should not be inhibited between the external surface of the stores showing the individuals. ii) to assign numerical values of compliance uses the concept of Pareto dominance. iii) the Pareto optimal surface characteristics of the clustering technique does not use without reducing the number of individuals that are stored externally. SPEA vary from four aspects: First, join the above three properties in only one algorithm. Second, compile a list of the suitability of the individual, determines not only the individuals stored in the external cluster. Abiding raided each other is unimportant. Third, all members of the external cluster selection string. And at last a
new Pareto-based Population niching method that proposes to maintain the diversity [15].

At the SPEA eligibility assignment, density estimation and potential weaknesses of archival reduction be healed again, Zitzler et. al. [16], and in 2001 proposed SPEA2.

Assignment Eligibility: Individuals repressed by members of the same archive, has identical fitness values. This means that the archive has a single individual in the case of the entire population members, each pressing down on or not down is to have the same degree of effectiveness of selection. As a result, the selection pressure is greatly reduced. In this case, SPEA works as random search algorithm.

Intensity approximation: If the existing population identical to many individuals, not dominate to each other, the dominance partial ordering relation defined on the basis of little or no information is obtained. Intensity information; search direction to be used more effectively to represent.

Archive Reduction: At SPEA used classification technique, despite reduction inhibited without impairing the characteristics of the cluster, individuals may lose discarded. However, these solutions to obtain a good distribution of solutions to be inhibited is better to keep the archive. This is carried out in SPEA2 [16].

Appointment of Compliance: SPEA is different from an individual only by members of the same archive to avoid pushing them to have identical values of compliance, both printed and taken into account by all individuals should not be inhibited. In each individual of the archive and in the population, the number of individual units of a repressing solution $S(i)$ is assigned a value of force.

$$S(i) = |\{ j \mid j \in P_t + P_t \wedge i \succ j \}|$$  \hspace{1cm} (7)

Where, $|.|$ Notation indicates that a cluster the most important-, +, and the combination of multi-cluster Pareto superiority symbol relates to the relationship. $I$ is calculated on the basis of the value of $S$ member compliance with the raw $R$. members is the sum of $S$ values of individuals $j$ suppresses $i$ returns the value of raw fitness.

$$R(i) = \sum_{j \in P_t + P_t \wedge j \succ i} S(j)$$  \hspace{1cm} (8)

Raw fitness assignment, a kind of based on the concept of Pareto superiority niching provide a mechanism, although many individuals may fail to suppress each other. Therefore, to distinguish between individuals with identical raw fitness values to an additional density information is used. $K$ as the density estimation inverse distance to the nearest neighbor is taken. Clearly, for each individual $i$, $j$ archive of all population members, and the distances are calculated and stored in a list. After listing in ascending order of this list, $k$. The distance between the elements, and this are indicated by. Generally, $k$, is taken as the square root of the size of the sample space.

$$k = \sqrt{N + N}$$  \hspace{1cm} (9)

Then, density is calculated as follows:

$$D(i) = \frac{1}{\sigma_i^k + 2}$$  \hspace{1cm} (10)

Here, the denominator is greater than zero and $D<1$ to 2 is added. I worked as a member of the fitness value of the fitness value of intensity is collected:

$$F(i) = R(i) + D(i)$$  \hspace{1cm} (11)

Environmental Choice: The first step in the process of environmental selection, all of which is lower than 1 compliance archives and inhibited population of individuals copy the next generation.
\begin{equation}
\mathbb{P}_{t+1} = \{i \mid i \in P_t + \mathbb{P}_t \land F(i) < 1\}
\end{equation}

(12)

If the size of the archive is fully inhibited the number of individuals are in accord with the surface of environmental selection process is completed. Otherwise, it should be the size of the archive is either small or large. Archive size is small, the previous archive should not be inhibited, and the best individual in the population are copied to the new archive. Archive size exceeds the number of individual cases as the auction is done iteratively until the size of the repository. To decide which individuals should be removed to apply the procedure of reduction to an archive.

\begin{equation}
i \leq j : \forall 0 < k < |\mathbb{P}_{t+1}| : s^i_k = s^j_k \lor
30 < k < |\mathbb{P}_{t+1}| : \left(\forall 0 < l < k : d^i_l = s^j_l\right) \land s^i_k < s^j_k
\end{equation}

(13)

Where \( s^i_k \) member is selected \( i \in \mathbb{P}_{t+1} \) at k. shows the nearest neighbor distance. Here, each step, the minimum distance to the others with a selected individual. If you have a minimum distance of a few individuals, if the latter is preferable, and the next neighbor distance, looking purged.

**Minimum Error Function:** Genetic algorithm to solve a problem with some of the important elements of the encoding, selection, crossover and mutation. One of the factors affecting the performance of GA chromosomes encoding model. In the study, all the points one bits (genes) will be represented by a chromosome is made based on the method of binary encoding. Therefore, \( N \) is the number of points of chromosome length. Fig. 2 shows a chromosomal formed by accepting the \( N = 10 \). In this setting, if the chromosome gene of "1" is selected as a gene that corresponds to one of the knots. Otherwise, the knot sequence of points taken into consideration. Initial population is created once and then starts to evolution. Blessings of individuals by finding the values of genetic algorithm tries to determine eligibility. Considered to be good again, according to the size of reproductive individuals of Conformity, crossover and mutation operators are selected. Points of the curve produced by the estimated objective function, do the minimum amount of the error between the given points as a function of the Euclidean error for the b u used for study.

\begin{equation}
|S_i - F_i| = \sqrt{(x_i - x_{i-1})^2 + (y_i - y_{i-1})^2}
\end{equation}

(14)

The total error, the equation (6) is obtained by using. According to all of SPEA2 algorithm is proposed for the prediction of B-spline curve can be given as follows.

**Step 1.1:** Generate initial population: N number of chromosomes is produced. Given the number of genes on chromosome must equal the number of knots. Each point corresponds to a knot in a gene on chromosome. Start and finish as the one assigned to the gene.

**Step 2:** Create Archive in the empty set.

**Step 3:** Environmental selection: all the individuals in the population and should not be inhibited in the archive which will be used in the next iteration Archive (Archive_t +1) set will be copied. Reduction in the size of the operation overflows Archive_t +1 is applied. Population size cannot be reached Archive_t +1, Archive_t copy and archive the best suppressed individuals.

**Step 4:** Termination: If \( t \geq T \) is the end. A set of individuals was inhibited in Archive_t +1 Copy, End.

**Step 5:** Mating selection: binary tournament selection on the mating pool to fill Archive_t +1.

**Step 6:** Change: Apply mating pool, crossover and mutation operations. Conclusion population Copy populasyon_t +1. Increment generation counter \( t = t +1 \). Go to Step 2.

## 5 Experimental Results

That the proposed estimation algorithm SPEA2-based B-spline curve to assess knots, Fig.3 shows an example of a B-spline curve parameterization. Fig.3. A 2D data shown in (200 units) to find a clean 10% noise added to the uniform distribution in the data. The bill is modeled curve, with 11 knot vector and control points of 7 cubic B-spline curve is a non-uniform. The Figure shown at the control polygon. Chromosomes in each generation is given the best results with precision. Effective value of the modeled curve and a point cloud (effective value-RMS) error is given in Table 1.
Fig. 3 a) 3-D data points form of B-spline (b) after 300 generation, parameterized B-spline curve. Table 1 The example shown in Fig. 3 a B-spline curve fitted to the RMS values for the Genetic Algorithm.

<table>
<thead>
<tr>
<th>Generation</th>
<th>The best RMS (GA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initially</td>
<td>390.57</td>
</tr>
<tr>
<td>10</td>
<td>371.14</td>
</tr>
<tr>
<td>25</td>
<td>364.26</td>
</tr>
<tr>
<td>50</td>
<td>358.41</td>
</tr>
<tr>
<td>100%</td>
<td>334.03</td>
</tr>
<tr>
<td>200</td>
<td>327.21</td>
</tr>
<tr>
<td>300</td>
<td>316.12</td>
</tr>
</tbody>
</table>

Initial population was fed up to 300 generations. Fig. 3b 300 is the next generation convergent pattern. The curve is now better suited to the data points. Generations of compliance with regulations is increasing with increasing (decreasing error). The slope of the curve shows that the probability of convergence to the next generations still. The values of parameters used in the experiment is as follows: the population size from 20 individuals, chromosomal length 200, the number of diversity 6 subjects (30%). To see the convergence speed of the proposed GA-based algorithm, some generations in the process of education programs are the outputs. This is according to the outcomes of individuals from populations that generation, maximum, minimum and average RMS error values are given in Table 2. SPEA2 approach recommended by the convergence of all generations in the chart presented in Fig. 4.

Table 2. the RMS statistics of the SPEA2 Optimization shown in Fig. 3

<table>
<thead>
<tr>
<th></th>
<th>Maximum</th>
<th>Minimum</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initially</td>
<td>3934</td>
<td>392</td>
<td>392</td>
</tr>
<tr>
<td>10</td>
<td>3820</td>
<td>376</td>
<td>372</td>
</tr>
<tr>
<td>25</td>
<td>3789</td>
<td>367</td>
<td>366</td>
</tr>
<tr>
<td>50</td>
<td>3833</td>
<td>363</td>
<td>358</td>
</tr>
<tr>
<td>100%</td>
<td>4076</td>
<td>640</td>
<td>370</td>
</tr>
<tr>
<td>200</td>
<td>13 122</td>
<td>4556</td>
<td>2278</td>
</tr>
<tr>
<td>300</td>
<td>3812</td>
<td>351</td>
<td>349</td>
</tr>
</tbody>
</table>

Fig. 4 SPEA2 parameter optimization according to Generations.

6 Conclusion

Given any number of irregular points, the large number of curve can be produced. B-spline curves offer flexibility and modification success to the users, thus they are type of industry-standard curve. For irregular data points, finding B-spline curve with the lowest error value and the number of knots is a multi-objective Combinatorial optimization. Single-objective optimization, an objective decision that is best for a single design or tried to be achieved. This is usually based on minimization or maximization problem maximum presence of the global minimum or global. Very different from single-objective optimization of multi-objective optimization. In the real world, almost all of the designs or multiple problems and requires optimization of conflicting objectives simultaneously. The optimal solution search space, quenched by any other solution is the solution. The optimal solution is called Pareto-optimal solution. This paper describes a method of meta-heuristic genetic algorithm for the Strength Pareto Evolutionary Algorithm 2 with a new variant of the B-spline curve has been sought for the problem. In general, the data fitting problem is difficult to find global optima (with a large number of local optima is continuous nonlinear and multivariate optimization problem), to overcome this difficulty paper translated and translated into the original problem of discrete combinatorial optimization.
problem, which solves the problem with Strength Pareto Evolutionary Algorithm 2 strategy is used. Algorithms can define the best spin control points or knots and for optimal parametric curve that represents a powerful tool, so all the power available, the B-spline. Uses Euclidean distance as a function of conformity with the simplicity of the model and candidate models are offset by the loyalty of the data model can be automatically selected best model. Smoothness factor and a well-chosen initial locations of knots don't needed, such as any subjective Measures. Paper shows the great potential of this approach. The next study will be taken from the literature VEGA, MOGA, NPGA other multi-objective optimization algorithms such as the genetic variant of performance evaluations and comparisons can be expanded with the implementation of the same problem. Also increased the number of test problems and the objective functions. In this case the algorithm convergence speed of the Pareto surface can be added to monitor and evaluate performance. Study the advantages and drawbacks of algorithms taking into consideration the principles set out in a certain period, the number of generation, distribution and diversity increased or decreased according to the Pareto surface crossover rate, mutation rate, the archive size, for the run-time parameters such as clustering techniques and distance metrics has the ability to Adjuvant A new algorithm can be improved. Thus the convergence of the Pareto optimal surface, showing faster and more reliably developed a new technique.

**Acknowledgment.** This work is partially supported by Turkish Scientific and Technological Research Council (TÜBİTAK) under Grant No110E016.

**References:**


[15] Zitzler E. 1999. Evolutionary algorithms for multiobjective optimization: Methods and applications, Doctoral dissertation ETH 13398, Swiss Federal Institute of Technology (ETH), Zurich, Switzerland