İstanbul Üniversitesi İşletme Fakültesi Dergisi
Istanbul University Journal of the School of Business Administration
Cilt/Vol:41, Sayı/No:1, 2012, 47-61
ISSN: 1303-1732 – www.ifdergisi.org © 2012

# Genetic algorithm for distance balancing in set partitioning problems

**Serap Kiremitci**[1]
*The School of Transport and Logistics*
*Istanbul University, Istanbul, Turkey*

**İbrahim Zeki Akyurt**[2]
*Department of Production Management*
*Faculty of Business Administration*
*Istanbul University, Istanbul, Turkey*

**Abstract**

In this study balancing is taken into consideration in the formation of groups according to the total travel distance as a set partitioning problem (SPP). Fitness functions that can test imbalance are proposed and mathematical models including these fitness functions are presented. In order to make balanced groupings, four different fitness functions are used. The first model aims to minimize total travel distance. The other two models are used for balancing and the last one constitutes a precedent as a multi-objective decision making problem. Genetic Algorithms (GAs) which is a meta-heuristic technique is used for the solution of the proposed models. Data is taken from the study of Akyurt et al. [1] and is used to balance groups in Football Leagues. Different groups are formed according to these models; effects of the results are examined among themselves and compared with the current situation. Additionally, all results are displayed on the maps.

***Keywords:*** *Balancing, Set Partitioning Problem, Genetic Algorithms, Football*

***Küme bölme problemlerinde genetik algoritma ile mesafe dengelenmesi***

***Özet***

Bu çalışmada Set Partitioning Problemlerinde (SPP) mesafeye göre gruplandırma yapılırken denge unsuru göz önünde bulundurulmuştur. Dengesizliği test edebilecek *fitness function*lar önerilerek bu fonksiyonları içeren matematiksel modeller kurulmuştur. Dengeli gruplandırmayı yapabilmek için 4 farklı *fitness function* kullanılmıştır. İlk model toplam mesafeyi minimum kılmayı hedeflemiştir. Diğer iki model dengeleme için kullanılmıştır, son model ise çok amaçlı karar verme problemine örnek oluşturmaktadır. Modeller bir meta-sezgisel yöntem olan Genetik Algoritmalar (GAs) ile çözülmüştür. Çalışmada, kullanılan veriler Akyurt vd. [1] yaptıkları çalışmadan alınmış ve Futbol Liglerinde dengeli gruplama üzerine uygulanmıştır. Tüm bu modellere göre farklı gruplandırmalar yapılmış; sonuçların kendi aralarında etkileri incelenmiş ve mevcut durum ile ayrı ayrı karşılaştırılmıştır. Ayrıca tüm sonuçlar harita üzerinde gösterilmiştir.

***Anahtar Sözcükler:*** *Dengeleme, Küme Bölme Problemleri, Genetik Algoritma, Futbol*

## *1. Introduction*

The primary objective of this paper is obtaining a balanced solution among sets and elements of sets as a set partitioning problem with Genetic Algorithms. For this purpose, imbalance ratios are determined and enclosed in a mathematical model. The mathematical model is developed in order to balance the groups and their elements which resulted from the SPP.

[1] *serapy@istanbul.edu.tr (S. Kiremitci)*

[2] *akyurt@istanbul.edu.tr (İ.Z. Akyurt)*

The proposed mathematical model is used for the determination of team members of each group in the Turkish Football Federation Third League Classification Groups. The first objective function aims to determine the teams for each group for the purpose of minimizing the total travel distance taken by the teams. Accordingly, the second and third objective functions are concerned with a balancing notion. While the second one addresses balancing between groups, the third one deals with a balancing among teams. The last objective function constitutes a precedent as a multi-objective decision making problem, *targeting the optimization of both the first and second equations together.* The goal of the last function is first tominimize the total travel distance and secondly to obtain a balance between groups. The Pareto method is used to balance these two values in this problem.

The set partitioning problem has been studied extensively over the years because of its many important applications [2]. As SPP has wide applicability in many real–life problems, it has both exact and heuristics methods for its solution. There are different methods among the exact methods. The most classic one is to solve SPP with linear programming. It is noted that obtaining the solution is very easy in small size instances with this one [3, 4, 5]. The other method among the exact ones, is the branch and bound method. They [6, 7, 8, 9, 10] used branch-and-bound methods to successfully solve several problems from airline crew scheduling and bus routing. Moreover the Lagrangian dual approach [11], the Branch and Cut Method [12], the Branch and Price Method [13] and the Column Generation Method [14] can be seen for the solution of SPP in literature. Fisher and Kedia [15] and Chan and Yano [16] proposed branch-and-bound algorithms that used a heuristic solution for dual linear relaxation programming in order to obtain bounds. Grossman and Wool [17] compared nine different algorithms.

The problem is NP-Hard that obtain results for large size problems with standard methods which are challenging and complicated, therefore heuristic algorithms have been developed. There are many papers concerning heuristic methods for the solution of SPP [18, 19, 20, 21, 22, 23, 24, 25, 26]. Genetic Algorithms (GAs) and a Tabu Search are the heuristic methods used. In references [27, 28, 29, 30, 31, 32], genetic algorithms are used for the solution of SPP and also usually for the applications of airline crew scheduling. The Tabu search method has also been used to solve SPP [33, 34].

On the other hand when we searched literature about the formation of groups in sports leagues, Güngör and Küçüksille [35] have used SPP in Football Leagues. They have separated 51 teams into three groups in the Turkish Football League B Category with a GAs based approach. In separation of the groups, distances between the teams are minimized. Akyurt et al. [1] have studied a similar problem and have achieved approximately 13% percent improvement in total travel distance. This paper extends the study of Akyurt et al. by adding a balancing notion. Relative imbalance ratios, similar to the ones used in Keskintürk et al. [36], are computed for the solution of distance balancing. Keskintürk et al. [36] presents a mathematical model for a parallel- machine problem with sequence-dependent setups where the goal is to minimize total relative imbalance. They used an ant colony optimization algorithm and a genetic algorithm for this purpose.

In this paper the proposed model for distance balancing of the set partitioning problem (DBSPP) is used in the Turkish Football Federation (TFF) Third League Classification Groups, teams during the 2009-2010 season are dealt with. There are 53 teams in 5 separate groups. The first four groups are composed of 11 teams and the last group is composed of 9 teams. During the season two matches are played between each of the teams and the other competitors in the same group. In the study, comparison of the groups determined by the Turkish Football Federation (TFF) and by the proposed genetic algorithm was made according to some criteria that includes total distance taken by the teams during the whole season, balance between groups and balance between teams.

While calculating the total travel distance matrices, distances between each team were obtained from Google Maps.

The paper is organized as follows: In the next section, after a brief description of Set Partitioning Problems is given, problem definition and balancing notion are also provided. In section 3, developed Genetic Algorithms to solve this problem is presented. Results of proposed algorithm on TFF Third League are discussed in Section 4. Some concluding remarks and perspectives for future research are presented in Section 5.

## *2. Problem Definition*

The Set Partitioning Problem (SPP) is known to be NP-hard and it can be used to model many important real-world decision problems [37] including those involving railroad crew scheduling, truck deliveries, airline crew scheduling, tanker routing, information retrieval, switching circuit design, stock cutting, assembly line balancing, capital equipment decisions, location of offshore drilling platforms, some other facility location problems, and political districting [3].

Set Partitioning problem can be formulated as follows:

$$Min \sum_{j}^{N} c_j x_j$$

$$Ax = e$$

$$x_j = \{0,1\}$$

A=$(a_{ij})$ is an m*n matrix which consists of 0 and 1 (elements). c is an arbitrary n-vector, e=(1,…,1) is an m-vector, and N={1,...,n}. If the rows of A are associated with the elements of the set M={1,...,m} and each column $a_j$ of A with subset of $M_j$ of those i € M such that $a_{ij}$=1, then SPP is the problem of finding a minimum-weight family of subsets $M_j$, j ∈ N, which is a partition of M, each subset being weighted with $c_j$.

In this study, teams' clustering is made in order to form groups in a regional football league. Assignment of teams to groups is made in such a way that the total travel distance will be minimized. This problem is modeled as a set partitioning problem. The mathematical models are modified according to different fitness functions. Since our goal is to obtain a balanced solution between teams and groups, we call this problem a distance balancing set partitioning problem (DBSSP). The problem notation is given as follows:

| | | |
|---|---|---|
| $N$ | : | Number of Teams |
| $n$ | : | Team index *n*=1,..,N |
| $K$ | : | Number of Groups |
| *k* | : | Group index *k*=1,.., *K* |
| $V$ | : | Set of Teams |
| $T_k$ | : | Number of teams in group *k* |
| $S_k$ | : | Set of teams in group $k,\ S \le V$ |
| $c_{(x_i x_j)}$ | : | Distance between team $x_i$ and $x_j$, $\forall\,(i,j) \in V$ |
| $\bar{c}_k$ | : | Average travel distance taken by a team in group k in a Season |
| $\bar{\bar{c}}$ | : | Average distance taken by a team in a season |
| $d_i$ | : | Total distance taken by team i in a season |
| $x_i$ | : | Problem variables, $\forall\, i \in V, x_i = [1, N]$ |

In the study of Akyurt et al., the goal is to determine the teams of each group for the purpose of minimizing the total travel distance taken by the teams of the Turkish Football Federation (TFF). Their integer formulation of the Mathematical Model (Problem) is given;

$$\boldsymbol{min} \sum_{\boldsymbol{k=1}}^{\boldsymbol{K}} \sum_{\boldsymbol{i \in S_k}} \sum_{\boldsymbol{j \in S_k}} \boldsymbol{c_{(x_i,x_j)}}, \quad \boldsymbol{i \neq j} \qquad 1$$

$$|\boldsymbol{S_k}| = \boldsymbol{T_k}, \quad \boldsymbol{k = 1, \dots, K}$$

$$\boldsymbol{x_i} = [\boldsymbol{1}, \boldsymbol{N}] \quad \text{and integer}$$

The objective function minimizes total travel distance (the sum of all travel distances of all teams in all k groups during one season). As in this league all the teams play matches twice, one in their home and the other in an opponent team's home, with all the other teams in the same group. This equation minimizes total travel distance but does not deal with balancing. First, balancing between groups is taken into consideration. Balancing between groups means to minimize the difference of the total travel distance of the groups. As the numbers of teams of the groups are different, it will not be meaningful to get closer to the total distance of the groups. Therefore, $\bar{c}_k$ which denotes average total distance of a team in group k, has been found for all groups.

$$\boldsymbol{\bar{c}_k} = \frac{\sum_{\boldsymbol{i \in S_k}} \sum_{\boldsymbol{j \in S_k}} \boldsymbol{c_{(x_i,x_j)}}}{\boldsymbol{T_k}}, \quad \boldsymbol{k = 1, \dots, K} \quad \boldsymbol{i \neq j}$$

$\bar{\bar{c}}$ is the average total travel distance of all groups (the average total travel distance of teams in a season).

$$\bar{\bar{c}} = \frac{\sum_{k=1}^{K} \bar{c}_k}{K} = \frac{\sum_{k=1}^{K} \sum_{i \in S_k} \sum_{j \in S_k} c_{(x_i,x_j)}}{N},$$

When the sum of the differences of these two parameters for each of the groups is divided by the number of groups, the average deviation of the average distance of groups has been calculated and called the```` Relative Imbalance Ratio of Groups (RIRG).

$$\boldsymbol{RIRG} = \sum_{\boldsymbol{k=1}}^{\boldsymbol{K}} |\boldsymbol{\bar{c}_k} - \boldsymbol{\bar{\bar{c}}}| \Big/ \boldsymbol{K}$$

Accordingly, balancing between the groups can be obtained by the minimization of RIRG. Second objective function enables this and can be seen below.

$$\boldsymbol{min} \sum_{\boldsymbol{k=1}}^{\boldsymbol{K}} |\boldsymbol{\bar{c}_k} - \boldsymbol{\bar{\bar{c}}}| \Big/ \boldsymbol{K} \qquad 2$$

The third objective function brings fairness between teams irrespective of which group the team takes place. $d_i$ which denotes the total travel distance of team i is formulated as follows:

$$\boldsymbol{d_i = \sum_{j \in S_k} c(x_i, x_j) \quad k = 1, \dots, K \quad i \in S_k \; i \neq j}$$

When the sum of the differences of all teams' total travel distance ($d_i$) from average total distance ($\bar{\bar{c}}$) is divided by number of teams, the Average Deviation of the Team's Distance has been found and called Relative Imbalance Ratio of Teams (RIRT).

$$RIRT = \sum_{i=1}^{N} |d_i - \bar{\bar{c}}| \Big/ N$$

Accordingly, balancing between the teams can be obtained by the minimization of RIRT. The third objective function enables this and formulated below.

$$\boldsymbol{min \sum_{i=1}^{N} |d_i - \bar{\bar{c}}| \Big/ N} \qquad 3$$

In this study, the grouping of teams has been emphasized finally in order to minimize not only the total travel distance but also the imbalance between groups. To do so, the first and second objective functions are combined in one equation. But these two goals must have similar importance on results so two coefficients are added into the equation which makes these two goals equivalent. These two coefficients are found by using an approximate Pareto. Consequently, the fourth objective function which minimizes the total travel distance as well as the relative imbalance ratio of the groups (RIRG) is formulated as follows:

$$\boldsymbol{min\left( KS_1 \sum_{k=1}^{K} \sum_{i \in S_k} \sum_{j \in S_k} c_{(x_i, x_j)}, + KS_2 \left( \sum_{k=1}^{K} |\bar{c}_k - \bar{\bar{c}}| \Big/ K \right) \right) \quad i \neq j} \qquad 4$$

Another objective function can be formulated which deals with the minimization of the total travel distance and the minimization of imbalance between teams like the fourth one. This will be similar with it but RIRT will be used instead of RIRG.

Although all these equations are used to obtain balanced groupings of football teams in this study; the same formulation can also be used for problems like machine scheduling, assembly line balancing, cell manufacturing planning in which balancing is an important issue. The coefficients used in the fourth objective function permit our formulation to be used for other problems.

### *3. Genetic Algorithms for DBSPP*

Genetic Algorithms (GAs) which is a population-based meta-heuristic technique, was developed by Holland [38, 39]. GAs evolve a population of individuals encoded as chromosomes by creating new generations of offspring through an iterative process until some convergence criteria are met [40]. Solution values (variables) are represented in the vectors named chromosomes. This representation may be binary coding or actual values can also be used. For a given number interval, initial solutions are generated with values determined randomly. In the problem, there are integer variables for each of the teams. These integer variables can take values between 1 and the total team number (N). According to this, there are variables as many as the team numbers and they can take integer values between [1, N] interval. In GAs, variables are defined in genes within chromosomes. Each of the chromosomes represents an alternative solution as shown in Figure 1.

| **Groups** | **Group 1** | | | | | | | | **Group 2** | | | | | | | | **Group 3** | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Team** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **1** | **2** | **3** | **4** |
| **Chromosome** | **13** | **18** | **2** | **9** | **5** | **15** | **10** | **20** | **4** | **12** | **8** | **16** | **1** | **19** | **14** | **6** | **11** | **17** | **3** | **7** |

First group's first team

**Figure 1 Chromosome Example**

A group of solutions are produced with the number of chromosomes (population size) and is called Initial Population. Then the quality of each solution candidate is evaluated according to the problem-specific fitness function. There are four different fitness functions in our model that is explained in the previous section. The creation of new generation belonging to the next iterations, are made through GA operators: selection, crossover and mutation.

As can be seen in the chromosome example, there are three groups which are composed, 8, 8 and 4 teams respectively, 20 teams in total. Permutation encoding is used here and each of the genes represents a team which takes place in the related group.

### *3.1. Selection*

Selection is an operator that chooses chromosomes from current population for the next generation. Chromosomes with a better fitness value have higher probability of being chosen for the next generation. In this paper, a roulette wheel selection process is used. A roulette wheel selection is used for selecting potentially good solutions for recombination. This fitness level is used to associate a probability of selection with each individual chromosome. This could be imagined similar to a roulette wheel in a casino. Usually a proportion of the wheel is assigned to each of the possible selections based on their fitness value. This could be achieved by dividing the fitness of a selection by the total fitness of all the chromosomes, thereby normalizing them to 1. Then, a random selection is made as the roulette wheel is rotated. Candidate solutions with a higher fitness value will be less likely to be eliminated but also there is still a chance that they may be.

### *3.2. Crossover and Mutation*

Crossover and mutation are two important operators that make changes in existing chromosomes in the search for better solutions. The aim of the crossover is to exchange information between the chromosomes, so it enables the creation of new and better individuals (chromosomes). In binary and real-value coding, the crossover step is taken by changing one side of the predetermined crossover point between two strings reciprocally. Thus, two new different individuals are obtained. In order to prevent unfeasible solutions in permutation encoding, one-point crossover is used as seen in Figure 2. Mutation consists of randomly modifying some gene(s) of a single individual at a time to further explore the solution space and ensure, or preserve, genetic diversity. The occurrence of mutation is generally associated with a low probability [41].

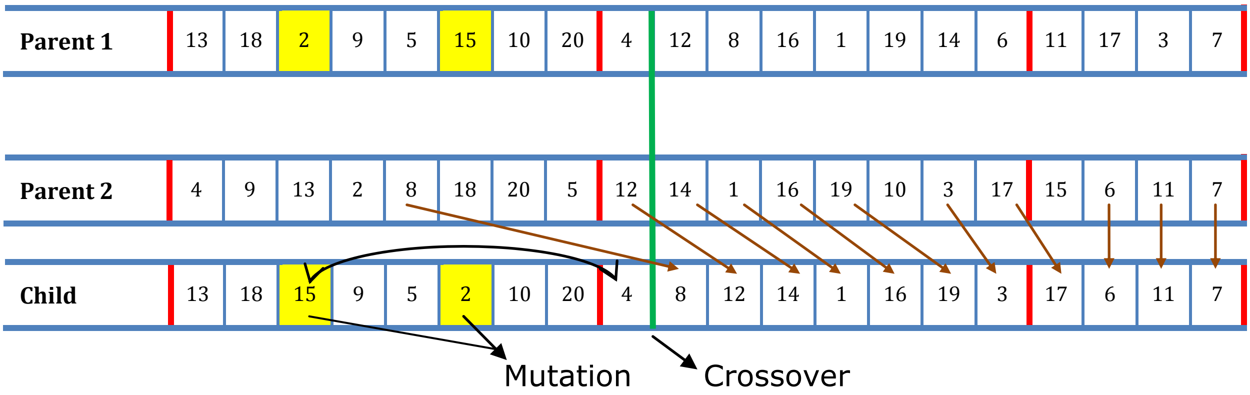


**Figure 2 Creation of a New Generation**

The steps of GAs are as follows:

**Step 1:** Generation of an initial population

**Step 2:** Evaluation of a fitness value

**Step 3:** Selection

**Step 4:** Crossover

**Step 5:** Mutation

**Step 6:** If stopping criterion is not achieved, go to the second step.

**Step 7:** Select the best solution as the result [41, 42].

This loop continues till getting the predetermined iteration number.

## *4. Results & Findings*

In this study data is taken from the of Akyurt et al. study [1]. Data is about the Turkish Football Federation Third League Classification Groups which have 53 teams in 5 separate groups during the 2009-2010 season. The first four groups are composed of 11 teams and the last group is composed of 9 teams. Distances between the teams are calculated by a developed program using coordinates as road distance in terms of kilometers which have been obtained from the Googlemaps.com website. In their study, the object is the determination of teams of each group for the purpose of minimizing the total travel distance [1]. For the proposed GA, the crossover rate was determined 0.9, the mutation rate was determined 0.001 and the population size was determined 20. The algorithm was coded with MATLAB R2009 and implemented in an Intel(R) Pentium(R) 4 CPU 3.20 GHz, 480 MB RAM configured PC for 1000 iterations, 100 times. Group based distributions of teams according to the TFF are shown in theTurkish Map in Figure 3.

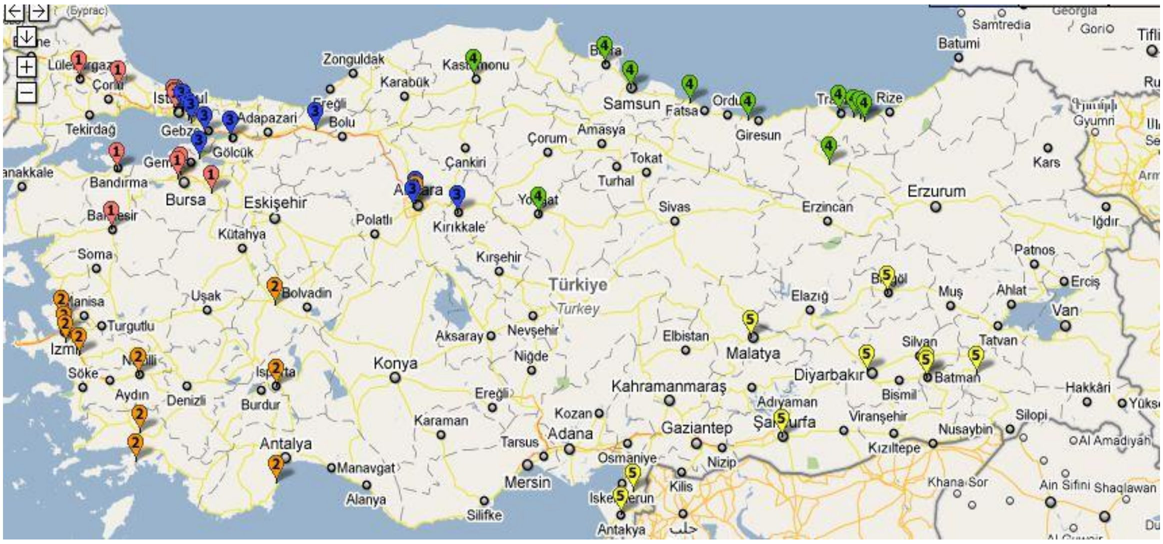


**Figure 3 TFF's Current Group Distribution on a Turkish Map**

Results obtained from equation 1 can be seen on Table 1 and Figure 4.

**Table 1 Results of Problem 1**

| | TFF | GAs (Mean) |
|---|---|---|
| **Total Distance** | **292164,4** | **253800,8** |
| **Imbalance Ratio of Groups (RIRG)** | 653,03 | 1371,49 |
| **Imbalance Ratio of Teams (RIRT)** | 1239,38 | 1561,85 |
| **Standard Deviation of Groups Distance** | 724,08 | 1736,73 |
| **Standard Deviation of Teams Distance** | 1700,97 | 2185,01 |
| **Variation Coefficient of Groups** | 0,131292 | 0,360503 |
| **Variation Coefficient of Teams** | 0,308565 | 0,456285 |
| **CPU** | - | 15,01 |

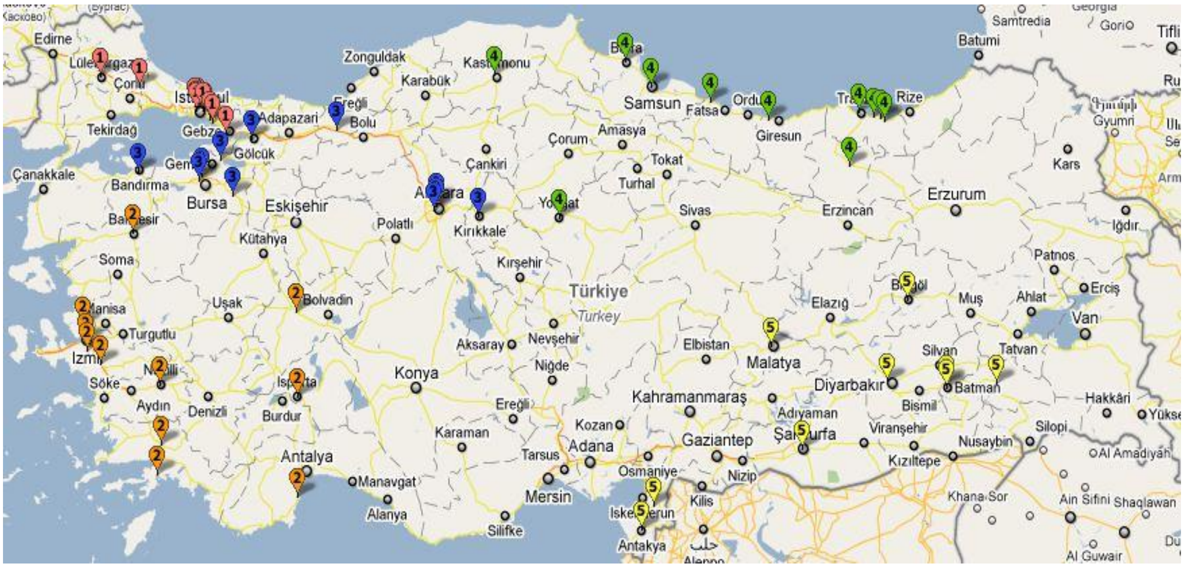


**Figure 4 Proposed GAs Group Distribution on a Turkish Map for Problem 1**

As shown in Tablo 1 the total travel distance taken by all 53 teams during the season has reduced from 292.164 km to 253.800 km, obtaining an approximate 13.1% improvement. This means 38363.6 km in terms of distance, but meanwhile RIRG has increased from 653 to 1371. This result reflects that balance between groups is not preserved. The same result can also be seen when we look at the standard deviation of groups. Results of Equation 2 which is used to obtain a balanced solution between groups can be seen at Table 2 and Figure 5.

**Table 2 Results of Problem 2**

| | TFF | GAs (Mean) |
|---|---|---|
| **Total Distance** | 292164,4 | 765348,6 |
| **Imbalance Ratio of Groups (RIRG)** | **653,03** | **1,69** |
| **Imbalance Ratio of Teams (RIRT)** | 1239,38 | 2667,57 |
| **Standard Deviation of Groups Distance** | 724,08 | 2,19 |
| **Standard Deviation of Teams Distance** | 1700,97 | 3281,95 |
| **Variation Coefficient of Groups** | 0,131292 | 0,000152 |
| **Variation Coefficient of Teams** | 0,308565 | 0,227274 |
| **CPU** | - | 1666,29 |


**Figure 5 Proposed GAs Group Distributions on a Turkish Map for Problem 2**

*(http://iubik.org/gunluk/baris/tff-3-lig-kademe-gruplari-2009-2010-kume-adaleti)*

According to Table 2, balance between groups has been obtained and RIRG has fallen to 1.69, but this time a significant increase in total distance can be seen. Results of Equation 3 can be seen in Table 3 and Figure 6 by the same way. The goal of Equation 3 is to obtain a balanced solution between teams. While obtaining the balance between teams, an improvement in terms of both total travel distance and balance between groups has been achieved.

**Table 3 Results of Problem 3**

| | TFF | GAs (Mean) |
|---|---|---|
| **Total Distance** | 292164,4 | 284481,8 |
| **Imbalance Ratio of Groups (RIRG)** | 653,03 | 506,65 |
| **Imbalance Ratio of Teams (RIRT)** | **1239,38** | **1080,55** |
| **Standard Deviation of Groups Distance** | 724,08 | 686,13 |
| **Standard Deviation of Teams Distance** | 1700,97 | 1525,89 |
| **Variation Coefficient of Groups** | 0,131292 | 0,127645 |
| **Variation Coefficient of Teams** | 0,308565 | 0,284280 |
| **CPU** | - | 294,46 |

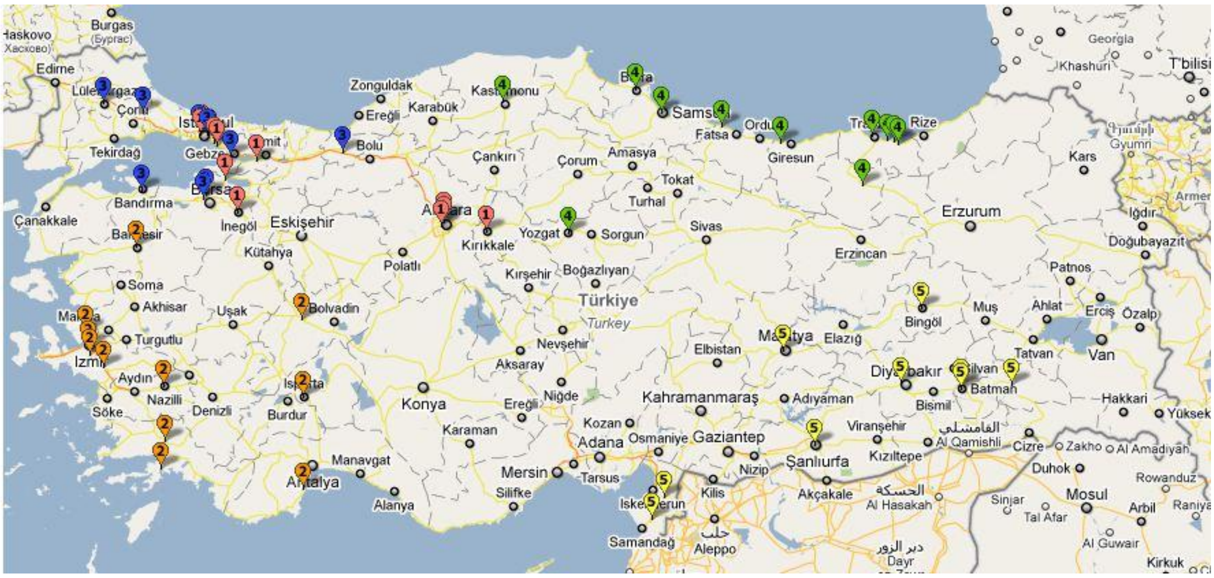


**Figure 6 Proposed GAs Group Distributions on a Turkish Map for Problem 3**

*(http://iubik.org/gunluk/baris/tff-3-lig-kademe-gruplari-2009-2010-takim-adaleti)*

Before computing Equation 4, coefficients have been evaluated and the best coefficient values are found with Pareto which is constructed from the results of trials. The value of KS1 has found 1 and the value of KS2 has found an interval of 40-44. Results from Equation 4 in which these coefficients are used, can be seen in Table 4 and Figure 7. Improvement is achieved not only in total travel distance but also in the balance between groups, but this time this improvement has increased the imbalance ratio of teams.

**Table 4 Results of Problem 4**

| | TFF | GAs (Mean) |
|---|---|---|
| **Total Distance** | **292164,4** | **277667,4** |
| **Imbalance Ratio of Groups (RIRG)** | **653,03** | **509,82** |
| **Imbalance Ratio of Teams (RIRT)** | 1239,38 | 1431,22 |
| **Standard Deviation of Groups Distance** | 724,08 | 682,10 |
| **Standard Deviation of Teams Distance** | 1700,97 | 2308,57 |
| **Variation Coefficient of Groups** | 0,131292 | 0,129889 |
| **Variation Coefficient of Teams** | 0,308565 | 0,440651 |
| **CPU** | - | 294,46 |

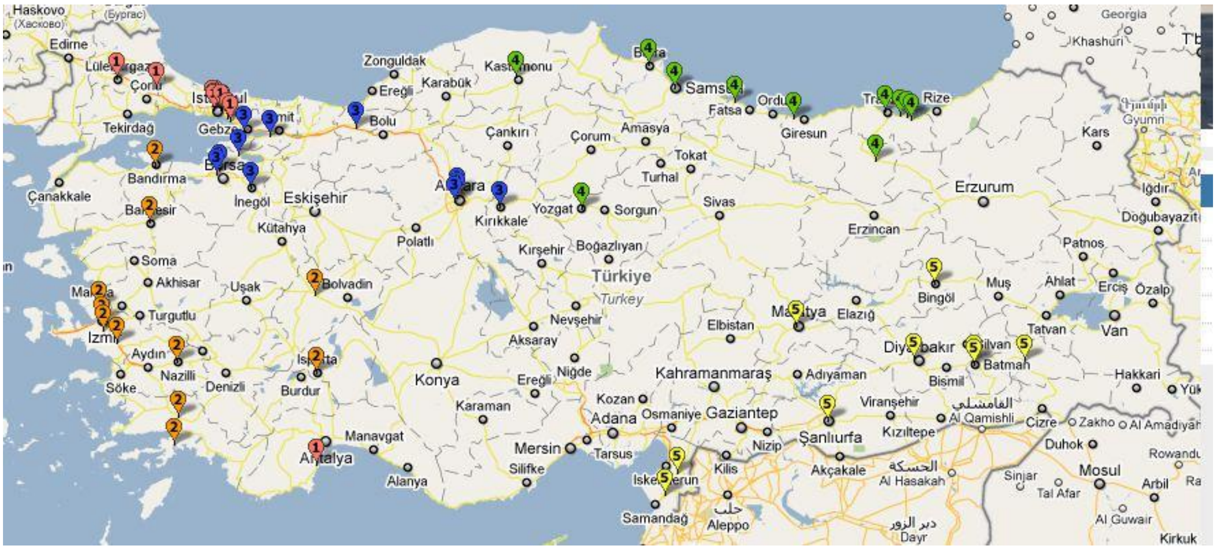


**Figure 7 Proposed GAs Group Distributions on a Turkish Map for Problem 3**

*(http://iubik.org/gunluk/baris/tff-3-lig-kademe-gruplari-2009-2010-toplam-mesafe-kume-adaleti)*

## *5. Conclusion*

In this paper, determination of the Turkish Football Federation's (TFF) 3rd League classification groups with genetic algorithm is presented as a set partitioning problem. The aim of the study is to provide an alternative solution with a scientific approach that deals with the minimization of both the total travel distance and imbalance, to TFF's current solution.

When the results are considered, the proposed GA's result is a good alternative to TFF's result in terms of distance balancing. According to the first fitness function which minimizes the total travel distance, the difference between the results in total distance is 38363.6 km, approximately 13.1%. With the second fitness function which minimizes RIRG, the value of RIRG has decreased from 653 to 1.69, showing the balance between groups has been achieved very well. But at the same time, the total travel distance increases approximately 2.5 times. According to the results of the third fitness function which minimizes RIRT for the purpose of balancing between teams, the value of RIRT has decreased from 1239 to 1080, making an improvement of approximately 13%. Also, a reduction has been observed in RIRG by 22% and in total distance by 2%. The last fitness function is an example for multi objective programming which seeks to minimize both the total travel distance and the relative imbalance ratio of groups. The total distance has been decreased to 277,667.4 km, providing a reduction of 5%, the value of RIRG has been decreased to 509, reflecting a 22% reduction. Meanwhile, there has been a 15% raise in the value of RIRT. When the results are examined, only the values at Table 3 reflect the results of the third objective function, which aims balancing between teams, has achieved improvement in all of the three criterias (total distance, RIRG, RIRT). This result cannot be valid for all problems as the goal is to obtain only the balance between teams. While achieving balance between teams, balancing of group's distances will also be provided. However, an increase in total distance can be observed. Thereby, considering total distance with the other imbalance ratio as in fitness function 4 may reveal better results in these types of problems. There are not significant differences between results because existing teams are distributed to specific regions and there are no teams in some regions as seen in Figure 3. If the algorithm will be applied to homogeneously distribute a bigger sample group, the differences between the results will become more meaningful.

The proposed model can also be applied to other areas such as machine scheduling, cell manufacturing planning and assembly line balancing problems.

As a result, it is considered that this improvement may enable serious benefits in economic, health and environment issues. Examining the figures, it can be seen that Proposed GAs forms better groups than TFF's and this enables an improvement in total distance and distance balancing.

### *References*


[1] I.Z. Akyurt, T. Keskintürk, B. Kiremitçi and S. Kiremitçi, Determination of Teams in Groups of Turkish Football Federation Third League Classification Groups by Genetic Algorithms*. VIII. International Logistics and Supply Chain Congress,* Istanbul, Turkey, (2010).

[2] E. Balas, and M.W. Padberg, Set Partitioning: A Survey. In N. Christofides, A. Mingozzi, P. Toth and C. Sandi (eds.), *Combinatorial Optimization*. John Wiley, 151–210, (1979).

[3] E. Balas, and M. Padberg, Set Partitioning: A Survey. *SIAM Review*, 18, 710–760, (1976).

[4] Gershkoff, Optimizing Flight Crew Schedules. *Interfaces*, 19, 29–43, (1989).

[5] R.E. Marsten, An Algorithm for Large Set Partitioning Problems. *Management Science*, 20, 774–787, (1974).

[6] R. Marsten, and F. Shepardson, Exact Solution of Crew Scheduling Problems Using the Set Partitioning Model: Recent Successful Applications. *Networks,* 11, 165–177, (1981).

[7] E. El-Darzi and G. Mitra, Solution of Set Covering and Set Partitioning Problems Using Assignment Relaxations. *Journal of Operations Research Society*, 483-493, (1992).

[8] E. El-Darzi and G. Mitra, Graph Theoretic Relaxations of Set Covering and Set Partitioning Problems. *European Journal of Operational Research,* 87, 109–121, (1995).

[9] J.E. Beasley and K. Jornsten, Enhancing an Algorithm for Set Covering Problems. *European Journal of Operational Research*, 58, 293–300, (1992).

[10] C.Mannino and A. Sassano, Solving Hard Set Covering Problems. *Operations Research Letters*, Volume 18, Issue 1, 1-5, (1995).

[11] D. Wedelin, An Algorithm for Large Scale 0-1 Integer Programing with Applications to Airline Crew Scheduling. *Annals of Operations Research,* 57, 283–301, (1995).

[12] K. Hoffman and M. Padberg, Solving airline crew scheduling problems by branch-and cut. *Management Science,* 39, 657–682, (1993).

[13] P. Avella, M. Boccia and A. Sforza, Solving a Fuel Delivery Problem by Heuristic and Exact Approaches. *European Journal of Operations Research*, 152, 170–179, (2004).

[14] I.D. Elhallaoui, Villeneuve, F. Soumis and G. Desaulniers, Dynamic Aggregation of Set-Partitioning Constraints in Column Generation. *Operations Research*, 53, 632–645, (2005).

[15] M.L. Fisher and P. Kedia, Optimal Solution of Set Covering/Partitioning Problems Using Dual Heuristics. *Management Science*, 36, 674–688, (1990).

[16] T.J. Chan and C.A. Yano, A Multiplier Adjustment Approach for the Set Partitioning Problem. *Operations Research,* 40, 40–47, (1992).

[17] T. Grossman and A. Wool, Computational Experience with Approximation Algorithms for the Set Covering Problem. *European Journal of Operational Research,* Volume 101, Issue 1, 81-92, (1997).

[18] E.K. Baker, L.D. Bodin, W.F. Finnegan and R.J. Ponder, Efficient Heuristic Solutions to an Airline Crew Scheduling Problem. *AIIE Transactions*, 11, 79–85, (1979).

[19] E. Balas and A. Ho, Set Covering Algorithms Using Cutting Planes, Heuristics and Subgradient Optimization: A Computational Study. *Mathematical Programming Study*, 12, 37–60, (1980).

[20] F.J. Vasko and G.R. Wilson, An Efficient Heuristic for Large Set Covering Problems. *Naval Research Logistics Quarterly*, 31, 163–171, (1984).

[21] M. Ball and A. Roberts, A Graph Partitioning Approach to Airline Crew Scheduling. *Transportation Science,* 17, 4–31, (1985).

[22] D.M. Ryan and J.C. Falkner, On the Integer Properties of Scheduling Set Partitioning Models. *European Journal of Operational Research*, 35, 422–456, (1988).

[23] J.E. Beasley, A Lagrangian Heuristic for Set-Covering Problems. *Naval Research Logistics,* 37, 151–164, (1990).

[24] L.W. Jacobs and M.J. Brusco, A Simulated Annealing-Based Heuristic for The Set-Covering Problem (Working Paper), *Operations Management and Information Systems Department,* Northern Illinois University, Dekalb, IL, March, (1993).

[25] S. Sen, Minimal Cost Set Covering Using Probabilistic Methods. *In: Proc. 1993 ACM/SIGAPP Symposium on Applied Computing,* 157–164, (1993).

[26] A. Atamtürk, G.L. Nemhauser and M.W.P. Savelsbergh, A Combined Lagrangian, Linear Programming and Implication Heuristic for Large-Scale Set Partitioning Problems. *Journal of Heuristics*, 1, 247–259, (1995).

[27] G. E. Liepins, M. R. Hilliard, M. Palmer, and M. Monow, Greedy Genetics, in Grefenstette J. J., editor, *Genetic Algorithm and Their Applications: Proceedings of the Second International Conference on Genetic Algorithms*, July, (1987).

[28] G. E. Liepins, M. R. Hilliard, J. Richardson, and M. Palmer, Genetic algorithms applications to set covering and traveling salesman problems, in Brown (ed), *Operations Research and Artificial Intelligence: The Integration of Problem-Solving Strategies*, Kluwer Academic Publishers, (1990).

[29] K.S. Al Sultan, M.F. Hussain and J.S. Nizami, A Genetic Algorithm for the Set Covering Problem. *Journal of the Operational Research Society*, 47, 702-709, (1996).

[30] D. Levine, A Parallel Genetic Algorithm For The Set Partitioning Problem. In: I.H. Osman and J.P. Kelly, Editors, *Meta-heuristics: theory and applications*, Kluwer Academic Publishers, Dordrecht 23–35, (1996).

[31] P.C. Chu and J.E. Beasley, Constraint Handling in Genetic Algorithms: The Set Partitioning Problem. *Journal of Heuristics*, 4, 323–357, (1998).

[32] J.E. Beasley, P.C. Chu, A Genetic Algorithm for the Set Covering Problem. *European Journal of Operational Research*, Volume 94, Issue 2, 392-404, (1996).

[33] J. Kelly and J. Xu, A Set Partitioning Based Heuristic for The Vehicle Routing Problem. *INFORMS Journal on Computing*, 11, 161–172, (1999).

[34] M. Lewis, G. Kochenberger and B. Alidae, A New Modeling and Solution Approach for the Set-Partitioning Problem. *Computers&Operations Research*, Volume 35, Issue 3, 807-813, (2008)

[35] İ. Güngör and E.U. Küçüksille, Küme Bölme Problemlerinin Genetik Algoritmayla Optimizasyonu: Türkiye İkinci Futbol Ligi Uygulması. *Review of Social Economics & Business Studies*, 5 (6), 381-394, (2005).

[36] T. Keskintürk, M.B. Yıldırım and M. Barut, An Ant Colony Optimization Algorithm for Load Balancing in Parallel Machines with Sequence-Dependent Setup Times. *Computers & Operations Research*, (article in Press).

[37] A. Joseph, A Concurrent Processing Framework for the Set Partitioning Problem. *Computers&Operations Research*, 29, 1375-139, (2002).

[38] D.E. Goldberg, *Genetic Algorithms in Search Optimization and Machine Learning*. Addison Wesley Publishing Company, USA, (1989).

[39] C.R. Reeves, *Modern Heuristic Techniques for Combinatorial Problems*. Mcgraw-Hill Book Company Inc., Europe, (1995).

[40] O. Braysy, M. Gendreau, Vehicle Routing Problem with Time Windows, Part II: Metaheuristics. *Transportation Science*, 39, 1, 119-139, (2005).

[41] İ.Z. Akyurt, T. Keskintürk, M.V. Arıkan, Determining The Parameters Of Periodic Optional Replenishment Model With Genetic Algorithm. *International Logistics And Supply Chain Congress*, Istanbul, Turkey, (2009).

[42] T. Keskintürk, Portföy Seçiminde Markowitz Modeli Için Yeni Bir Genetik Algoritma Yaklaşımı. *Yönetim*, 18 (56), 81, (2007).