Available online at www.elixirpublishers.com (Elixir International Journal)

Production Management

Elixir Prod. Mgmt. 67 (2014) 21716-21720

ISSN: 2229-712X

Primal-Dual Scatter Search with Tabu for Set Partitioning: Optimization of Football Leagues

Lutfu Sagbansua and Ibrahim Gungor
Department of Management, Turgut Ozal University.

ARTICLE INFO

Article history:
Received: 29 July 2013;
Received in revised form:
28 January 2014;
Accepted: 13 February 2014;

Keywords
Set Partitioning,
Scatter Search,
Tabu Search,
Heuristic,
Optimization,
Football Leagues.

ABSTRACT

Set partitioning problems are known to be NP-hard, thus it requires massive amounts of times and efforts to solve them using linear programming and traditional algorithms. This study proposes to use a scatter search algorithm coupled with a tabu search algorithm for such problems. The proposed algorithm is applied to the 3rd level football leagues in Turkey. 54 teams competing in the league are divided into three categories randomly by the Turkish Football Federation. The proposed algorithm in this study aims to set up these categories with the goal of minimizing the total amount of travelling, thus cost and time throughout the league. Experimental results show that the proposed algorithm reduces the total travelling by a significant amount of 56%.

Introduction

Dividing a set of units into a number of subsets, based on a determined objective can be approached as a Set Partitioning (SP) problem. The set partitioning problem arises when each unit must appear in exactly one subset, and all constraints are equality constraints. This type of problems has a large field of application including workforce planning (Yan and Tu, 2002; Yan and Chang, 2002; Sarin and Aggarval, 2002; Freling etal. 2001), vehicle routing (Chiang and Russell, 2002; Renaud and Boctor, 2004; Fagerholt, 1999; Kelly and Xu, 1999), planning for the best performance of various vehicles and machines (Chen and Lee, 2002; Fagerholt, 2001; Jiang, 2005), and alignment of sports leagues (Mitchell, 2003). However, solving these problems using linear programming techniques and traditional algorithms requires massive amounts of times and efforts. Thus, the search for more effective algorithms still continues.

Fugenschuh presents an integer programming model for the integrated optimization of bus schedules and school starting times, which is a single-depot vehicle scheduling problem with additional coupling constraints among the time windows. Set partitioning relaxation is applied to compute better lower bounds and, in combination with a primal construction heuristic, also better primal feasible solutions (Fugenschuh, 2011).

Linderoth and Lee suggest a parallel linear programming based heuristic approach for large scale set partitioning problems (Linderoth and Lee, 2001). Joseph proposes a concurrent processing framework to solve the large scale set partitioning problems by dividing it into smaller subgroups (Joseph, 2002). Barahona and Anbil use a volume algorithm and the dual simplex together (Barahona and Anbil, 2002). Klabjan suggests an algorithm for computing a subadditive dual function (Klabjan, 2004). Chu and Beasley apply well-known genetic algorithm in different ways to solve the SP problems of Lin, Kao, and Hsu (Chu and Beasley, 1998; Lin, Kao and Hsu, 1993). Combs combines tabu search and set partitioning approach for the crew scheduling problem (Combs, 2002).

The research at hand utilizes a scatter search algorithm combined with a tabu search component in the improvement phase to solve the SP problem introduced here. The application consists of grouping 54 teams competition in the 3rd level football league of Turkey. The format of the league is designed to have three groups consisting 18 teams in each. Since every team has to visit all other teams in the same group, an optimal solution must be found to minimize the total traveling time throughout the league, thus minimizing the overall cost and time required. Optimal solution can be found by setting up a 0-1 integer linear programming model and solving it to optimality (Mitchell, 2003).

Caballero etal. (2011) propose a hybrid heuristic procedure based on scatter search and tabu search for the problem of clustering objects to optimize multiple criteria with the goal of searching for good approximations of the efficient frontier for this class of problems and provide a means for improving decision making in multiple application areas.

The Mathematical Model for SP

The set partitioning problems can be formulated as follow (Van Krieken etal, 2003):

$$Min\ Z = \sum_{j=1}^{n} C_j X_j \tag{1}$$

subject to:

$$\sum_{j=1}^{n} a_{ij} X_j = 1 \qquad i = 1, \dots, m \tag{2}$$

$$X_j \in \{0,1\} \qquad j = 1, \dots, n \tag{3}$$

C_j in the objection function is the cost of j^{th} subset, thus $C_j \geq 0$.

Tele:
E-mail addresses: lsagbansua@turgutozal.edu.tr

$X_j = 1$, if the j^{th} subset is one of the subsets to be chosen for the optimal solution; otherwise, $X_j = 0$.

$a_{ij} \in \{0,1\}$. a_{ij} is a 0-1 matrix of mxn dimension. $a_{ij} = 1$, if the i^{th} unit is in the j^{th} subset; otherwise, $a_{ij} = 0$.

m , is the number of units to be divided into subsets.

n , is the number of subsets.

$Equation\ (2)$ ensures that each unit is placed in only one subset.

Primal-Dual Scatter Search-Tabu Algorithm for SP

Scatter Search (SS) is introduced by Glover initially for solving Integer Programming problems. However, the application range of SS has been expanded to cover combinatorial optimization and nonlinear optimization, in recent years.

SS derives its foundations from strategies originally proposed for combining decision rules and constraints (Glover et al., 2003). It operates on a set of solutions, called the reference set. The size of the reference set is usually no more than twenty (Glover et al., 2003). Two or more solutions of the reference set are chosen in a systematic rather than at random way to produce a new solution to update the reference set. When the reference set does not change and the algorithm reaches a specified iteration step, the solution is obtained (Zhang etal., 2011).

Tabu search (TS) algorithm applied in this study on the other hand, searches for the best solution starting with the solution provided by the SS during the solution combination phase. Each solution combination is subjected to an improvement phase equipped with a tabu search (TS) algorithm. Since the solution combination phase results with infeasible solution, improvement phase is designed to be able to deal with such solutions.

The improvement method involves shift and swap moves, first suggested by Martello and Toth (1981), on a given solution to reach another solution, with the goal of reducing the objective function value. If the new value is larger than the value of current solution, the move is not executed and the current solution is kept. Otherwise, the move is executed and the solution is updated. The first type of move, exchanges two teams in the same column of solution matrix, whereas a second type of move exchanges two teams located in different columns. At each iteration of the improvement, all possible shift and exchange moves are considered, and among those moves, the one that provides the best improvement in terms of the objective function value is selected as the next move (Sagbansua, 2009).

Improvement Algorithm

1. For each j^{th} team let

$$s_1(j) = \max_{i, i \neq i} \left\{ c_{i(j)j} - c_{ij} \right\}$$

with j_1 the maximizing team for $s_1(j)\ and\ i_1(j)$ the maximizing argument in the expression. (Assuming that the maximum over a null set is $-\infty$.)

2. For each pair of teams w and v let

$$s_2(v,w) = c_{i(v)v} + c_{i(w)w} - c_{i(w)v} + c_{i(v)w}$$
$$s_s(v,w) = \max_{v \neq w} \{s_2(v,w)\},$$

with maximizing arguments $v*\ and\ w*$.

3. If $s_1 = -\infty\ and\ s_2 = -\infty$ then STOP, as no more improvements can be found.

If $s_1 \geq s_2$, then shift $i_1(j_1)$ to the group j_1. Otherwise, if $s_2 > s_1$, then, exchange the teams $v*\ and\ w*$.

4. Go to step 1.

The resulting improved solutions are provided back to SS as candidates for Reference Set, during the Reference Set Update phase explained in the following section.

TS explores the solution space by moving at each iteration from a solution to the best available solution in its neighborhood. To avoid cycling, solutions possessing some attributes of recently explored ones are assigned the status of *tabu*. That is achieved using what is called short-term memory. Tabu moves are represented by attributes which are stored in a *tabu list*. The best move is chosen as the highest evaluation move in the neighborhood of the current solution in terms of objective function and tabu restrictions.

The SP problem studied in this research involves the third level Turkish football league consisting of 54 teams grouped by Turkish Football Federation (TFF) into three for the 2010-2011 football season. These teams are assigned a number to be represented by during the solution procedures. Naturally, each group consists of 18 teams which travel to each other for away games. Thus, the smaller the distances to each other, the smaller the total cost throughout the league. An mxm matrix generated to represent the distances between the teams. In this A matrix, $A[v][w]$ is the distance between the v^{th} and w^{th} teams. The distances are obtained from the General Directorate of Highways of Turkey.

Each solution visited by the algorithm is represented by a $3x18$ matrix *G*, where 3 is the number of subsets, and 18 is the number of teams in each subset.

Scatter Search Template

Suggested by Glover (1998), the following template has become the foundation of almost every scatter search algorithm available in the literature. The main idea behind using this template involves generating a diverse set of solutions along with a number of quality solutions and combining these solutions in the later stages while updating the reference set in order to keep diversification as well as the quality of reference set.

A Diversification Generation Method to generate a collection of diverse trial solutions, using an arbitrary trial solution (or seed solution) as an input.

An Improvement Method to transform a trial solution into one or more enhanced trial solutions. (Neither the input nor the output solutions are required to be feasible, though the output solutions will more usually be expected to be so. If no improvement of the input trial solution results, the “enhanced” solution is considered to be the same as the input solution.)

A Reference Set Update Method to build and maintain a reference set consisting of the *b* “best” solutions found (where the value of *b* is typically small, e.g., no more than 20), organized to provide efficient accessing by the other parts of the method. Solutions gain membership to the reference set according to their quality or diversity.

A Subset Generation Method to operate on the reference set, to produce a subset of its solutions as a basis for creating combined solutions.

A Solution Combination Method to transform a given subset of solutions produced by the Subset Generation Method into one or more combined solution vectors.

The Scatter Search Algorithm Description

The scatter search algorithm starts with a RefSet = $\emptyset$ (RefSet = $RefSet_1 \cup RefSet_2$). $RefSet_1$ contains three elements while $RefSet_2$ contains two. $RefSet_1$ is the subset where the elite solutions are added as an attempt to keep these solutions during the search process. Meanwhile, a set of diverse solutions are required in order to avoid getting trapped in a region. $RefSet_2$ is intended to serve this purpose by keeping diverse solutions to guide the search algorithm into new regions. The solutions obtained from the Diversification Generation Method are chosen to initialize this subset.

The algorithm starts with an initial solution which is chosen to be the current grouping scheme suggested by the TFF. The solution is improved using the Tabu algorithm introduced above. During the improvement iterations, elite solutions are added to the $RefSet_1$ if a solution improves the worst solution currently in this subset. In each iteration of the Improvement Phase, the solutions are checked to see whether they qualify to be included in the reference set. Improved solutions are evaluated in terms of their objective function values as well as a specified diversification criterion. Thus, the reference set is divided into two subsets: $RefSet_1$ containing a set of the best solutions in terms of the objective function value, and a $RefSet_2$ made up of diverse solutions. We consider a sequential evaluation process that checks whether the solution qualifies for $RefSet_1$, otherwise checks for $RefSet_2$. A solution qualifies for $RefSet_1$ if its objective function value is better than the objective function value of the worst element in this set, and in this case, the worst solution is replaced by the new one. Otherwise, we look at the diversity measure of the new solution. If the minimum distance of this solution to the current solutions in reference set is greater than the distance of any other solutions in the $RefSet_2$, then we replace this new solution with the worst solution in $RefSet_2$.

In the subset generation phase, two, three, and 4-elements subsets are generated by using the current reference set. Then, the Solution Combination Method is applied to find a new solution which is not guaranteed to be feasible. For instance, any team can be placed in more than one group. Thus, the improvement phase is designed to able to deal with such cases as well. The improvement phase follows the solution combination method in order to reach feasibility and further improve the objective value, consequently.

When we reach a maximum number of iterations, the best solution in terms of the objective function value is used to determine the new weights for the Surrogate Relaxation phase. After solving the Surrogate relaxed problem, the Improvement Method again comes into play, and the procedure mentioned above continues.

A description of the outline of the algorithm is provided below:

Initialization

1. Start with RefSet = $\emptyset$

For (Initial RefSet < RefSet)

2. Use Diversification Generation Method to generate a solution *x*. Apply the Improvement Method with TS feature to the solution to obtain improved solution *x**. Add the improved solution to RefSet.

End for

3. Order the solutions in RefSet according to their objective function value.

4. Take the first three solutions and add them to $RefSet_1$. Determine $RefSet_2$ based on the distance measures. Set RefSet = $RefSet_1 \cup RefSet_2$.

Primal-Daul

Do While (Max Iterations)

Do While (Max Tabu Iterations)

5. Apply Tabu search followed by improvement method to obtain a solution x_t.

If (x_t is not in RefSet and the objective function value of x_t is better than the objective function value of the worst solution in $RefSet_1$) **then**

Replace x_t with the worst solution in $RefSet_1$.

Else If (x_t is not in $RefSet_2$ and $d_{min}(x_t)$ is larger than $d_{min}(x)$ for a solution x in $RefSet_2$) **then**

Replace the worst solution currently in $RefSet_2$ with x_t.

Endif

End While

6. Order the solutions in $RefSet_1$ according to their objective function value.

7. Create subsets with the Subset Generation Method, using the solutions in RefSet.

Do While (Number of Subsets)

8. Apply the Solution Combination Method to obtain new solutions.

9. Apply Improvement Method with TS to obtain feasible solutions x_s.

If (x_t is not in RefSet and the objective function value of x_t is better than the objective function value of the worst solution in $RefSet_1$) **then**

Replace x_t with the worst solution in $RefSet_1$.

Apply Subset Generation Method

Else If (x_t is not in $RefSet_2$ and $d_{min}(x_t)$ is larger than $d_{min}(x)$ for a solution x in $RefSet_2$) **then**

Replace the worst solution currently in $RefSet_2$ with x_t.

Endif

10. Order the solutions in $RefSet_1$ according to their objective function value.

End While

End While

Experimental Results

The algorithm developed in this study is coded in C++ and run on a Pentium 1.73GHz processor, 0.99 GB of RAM.

The computational experience consists of 54 teams competing in the 3rd level national football league of Turkey in three groups for the 2010-2011 season. The SS algorithm utilized to minimize the total distance travelled by the teams during the season, uses the groups set up by the football federation as an initial solution. The distances between the team locations are obtained from the General Directorate of Highways of Turkey.

Considering each team in the league has 17 away games, 54 teams throughout the league have 918 away games totaling up to 726152 km (451209 miles) according to the groups set by the football federation. The solution above found by the SS algorithm on the other hand, requires 339589 km (211256 miles) which is a 56% reduction in the total traveling time, and the traveling costs consequently.

Concluding Remarks

The search for more efficient and suitable algorithms to solve NP-hard problems continues due to the insufficiency of linear programming techniques and traditional algorithms for such problems. Heuristic approach which emerges in near-optimal solutions very quickly has drawn a great deal of attention. Scatter Search introduced by Fred W. Glover is proven to be one of the most popular and effective metaheuristic methods.

In this research, we considered the multi-resource generalized assignment problem and proposed a new algorithm that consists of an effective use of the RAMP and Primal-Dual RAMP methods proposed in Rego (2004). It was confirmed through the computational results and the comparisons on the benchmark problem instances that these methods are very effective for the MRGAP which constitutes the first application of these methods in combinatorial optimization problems.

MRGAP is one of the combinatorial optimization problems which are known to be NP-hard.

This study develops a Scatter Search based algorithm to solve a Set Partitioning problem involving a national football league. The 54 teams competing in the league are divided into three subsets in a way to minimize the total distance travelled by the teams. The experimental results show that the proposed algorithm finds a solution reducing the total travel, hence the cost by 56% compared to the groups set up by the football federation.

The new groups are regional in nature. Thus, the structure of regional conferences strikes as a possible scenario that can be implemented. Such systems utilized by the National Football League (NFL) or National Basketball Association (NBA) can also be a model to minimize the costs while arranging the schedule in a way to add cross-conference games in certain time frames, as a way of diversification among the teams.

Future direction in this study is to further improve the Primal-Dual SST algorithm developed during this research. First tool in this direction is to use weight criteria in combining the elements in the subsets during the solution combination phase. This way, a more sophisticated and guided method will be utilized in this phase. The second tool is to improve the solution time required by the algorithm. The effectiveness of the modules used in the algorithm will be increased in order to save time, enabling more intensified search in the solution space.

References

BARAHONA, F., ANBIL, R. 2002. On Some Difficult Linear Programs Coming From Set Partitioning, Discrete Applied Mathematics, 118/1, 3-11.

CABALLERO, R. R., LAGUNA, M. M., MARTI, R. R., and MOLINA, J. J. (2011). Scatter tabu search for multiobjective clustering problems. Journal Of The Operational Research Society, 62(11), 2034-2046.

CHEN, Z.L., LEE, C.Y. 2002. Parallel Machine Scheduling with a Common Due Window, European Journal of Operational Research, 136/3, 512-527

CHIANG, W-C., RUSSELL, R. A.. 2002. Integrating Purchasing and Routing in a Propane Gas Supply Chain, European Journal of Operational Research, 154/3, 2004, 710-729

CHU, P.C., BEASLEY, J.E. 1998.Constraint Handling in Genetic Algorithms: The Set Partitioning Problem, Journal of Heuristics, 4/4,323-357.

COMBS, T.E. 2002. A Combined Adaptive Tabu Search and Set Partitioning Approach for the Crew Scheduling Problem with an Air Tanker Crew Application, Department of the Air Force Air University.

FAGERHOLT, K. 1999. Optimal Fleet Design in a Ship Routing Problem, International Transactions in Operational Research, 6/5, 453-464

FAGERHOLT, K. 2001. Ship Scheduling with Soft Time Windows: An Optimization Based Approach, European Journal of Operational Research, 131/3, 559-571

FRELING, R., LENTINK, R. M., ODIJK, M. A. 2001. Scheduling Train Crews: A Case Study For The Dutch Railways, In: Voß, S., Daduna, J.R. (Eds.), Computer-Aided Scheduling of Public Transport, Lecture Notes in Economics and Mathematical Systems, 505, Springer-Berlin, 153-165.

FUGENSCHUH, A. (2011). A set partitioning reformulation of a school bus scheduling problem. Journal Of Scheduling, 14(4), 307-318.

GLOVER, F. 1989. Tabu Search — Part I, ORSA Journal on Computing 1989 1: 3, 190-206.

GLOVER, F. 1990. Tabu Search — Part II, ORSA Journal on Computing 1990 2: 1, 4-32.

GLOVER, F., and M. LAGUNA. Tabu Search. Kluwer Academic Publishers, USA, 1997.

JIANG, M. 2005. UPS-k: A Set Partitioning Problem With Applications in UPS Pickup-Delivery System, Information Processing Letters, 93/4, 173-175.

Glover, F, M Laguna and R Marti. 2003. Scatter Search. In A Ghosh and S Tsutsui (eds.), Advances in Evolutionary Computing: Theory and Applications. New York: Springer-Verlag, pp. 519–537.

JOSEPH, A. 2002. A Concurrent Processing Framework For The Set Partitioning Problem, Computers & Operations Research, 29/10, 1375-1391.

KELLY, J. P., XU, J. 1999. A Set-Partitioning-Based Heuristic for the Vehicle Routing Problem, INFORMS Journal on Computing, 11/2, 161-172

KLABJAN, D. 2004. A Practical Algorithm for Computing a Subadditive Dual Function for Set Partitioning, Computational Optimization and Applications, 29/3, 347-368.

LÝN, FT., KAO, CY., HSU, CC. 1993. Applying The Genetic Approach To Simulated Annealing in Solving Some Np-Hard Problems, IEE Transactions On Systems Man And Cybernetics, 23/6, 1752-1767.

LINDEROTH, J. T., LEE, E. K., SAVELSBERGH, M. W.P. 2001. A Parallel, Linear Programming Based Heuristic for Large Scale Set Partitioning Problems, Informs Journal on Computing, 13/3, 191-209.

MARSEGUERRA, M., ZIO, E., PODOFILLINI, L. 2005. Multiobjective Spare Part Allocation By Means of Genetic Algorithms And Monte Carlo Simulation, Reliability Engineering and System Safety, 87/3, 325-335.

MARTELLO, S. and P. TOTH. 1981. An Algorithm for the Generalized Assignment Problem. *Operational Research, J.P. Brans, ed. North-Holland*, 589-603.

MAZZOLA, J., and S. WILCOX. 2001. Heuristics for the Multi-Resource Generalized Assignment Problem. Naval Research Logistics 48, 468-483.

MITCHELL, J. E. 2003. Realignment in the National Football League: Did They Do It Right?, Naval Research Logistics**,** 50/7, 683-701

REEVES, C. R. 1995. Modern Heuristic Techniques for Combinatorial Problems. McGraw-Hill International Limited, UK, 1995.

RENAUD, Jacques, BOCTOR, Fayez F., A Sweep-Based Algorithm For The Fleet Size And Mix Vehicle Routing Problem, European Journal of Operational Research, 140/3, 618-628.

SAGBANSUA, L. 2009. A New Algorithm for Minimum Cost MRGAP, European Journal of Economic and Political Studies, vol.2 no.2, pp.23-40.

SARIN, Subhash C., AGGARWAL, S. 2001. Modeling and Algorithmic Development of a Staff Scheduling Problem, European Journal of Operational Research, 128/3, 558-569.

VAN KRÝEKEN, M.G.C., FLEUREN, H., PEETERS, M.J.P. 2003. Problem Reduction in Set Partitioning Problems, Discussion Paper, Tilburg University.

YAN, S., CHANG, J-C. 2002. Airline Cockpit Crew Scheduling, European Journal of Operational Research, 136/3, 501-511.

YAN, S., TU, Y-P. 2002. A Network Model For Airline Cabin Crew Scheduling, European Journal of Operational Research, 140/3, 531-540.

ZHANG, J., TANG, J., FUNG, R. 2011. A Scatter Search for Multi-Depot Vehicle Routing Problem with Weight-Related Cost, Asia-Pacific Journal of Operational Research, 28/3, 323-348.