A Simulated Annealing Algorithm for Unsplittable Capacitated Network Design

M. Yaghini & M.R. Akhavan Kazemzadeh

Masoud Yaghini, School of Railway Engineering, Iran University of Science and Technology
Mohammad Rahim Akhavan, Department of Railway Transportation Engineering, Kermanshah University of Science and Technology

KEYWORDS
Network Design Problem, Unsplittable Multicommodity Capacitated Network Design Problem, Metaheuristic, Simulated Annealing

ABSTRACT
The Network Design Problem (NDP) is one of the important problems in combinatorial optimization. Among the network design problems, the Multicommodity Capacitated Network Design (MCND) problem has numerous applications in transportation, logistics, telecommunication, and production systems. The MCND problems with splittable flow variables are NP-hard, which means they require exponential time to be solved in optimality. With binary flow variables or unsplittable MCND, the complexity of the problem is increased significantly. With growing complexity and scale of real world capacitated network design applications, metaheuristics must be developed to solve these problems. This paper presents a simulated annealing approach with innovative representation and neighborhood structure for unsplittable MCND problem. The parameters of the proposed algorithms are tuned using Design of Experiments (DOE) method and the Design-Expert statistical software. The performance of the proposed algorithm is evaluated by solving instances with different dimensions from OR-Library. The results of the proposed algorithm are compared with the solutions of CPLEX solver. The results show that the proposed SA can find near optimal solution in much less time than exact algorithm.

1. Introduction
The network design problem is one of the important problems in combinatorial optimization. The objective of NDP is to find a minimum cost network on the available arcs for each commodity, which satisfy flows of commodities. Network design models have numerous applications in various fields such as transportation [38] [6], telecommunication [40] [20]; [41], and distribution planning [33].

One type of the network design models is splittable Multicommodity Capacitated Network Design (MCND) Problem. In this model, multiple commodities such as goods, data, people, etc., must be routed between different points of origin and destination on the arcs with limited capacity. In this problem, in addition to a unit cost (variable cost), a fixed cost is usually added due to the opening cost for the first time the arc is used. The MCND problem seeks a network with minimum cost which satisfies demands of commodities. This minimum cost is the sum of the fixed and variable cost [26]. In contrast with splittable MCND, flows of commodities cannot be separated on different routes in unsplittable MCND [38].
Network design problems can be easy stated but solving them is too difficult [2]. There are effective exact solution methods for uncapacitated network design problem. In exact algorithms, finding optimal solution is guaranteed.

Benders decomposition and branch-and-bound methods have been the two most effective methods for this type of problem [38]. The linear programming relaxations do not provide good lower bounds, and therefore simplex-based branch-and-bound methods are unlikely to succeed in solving even small instances. Indeed, most of the literature on exact approaches for the MCND problems and related problems uses Lagrangian relaxation [8]. Several surveys on network design models and their exact solution methods can be found in [38], [40], [2], [8], [42], and [17]. In addition, in [20], a good survey on models and solution methods for multi-commodity capacitated network design problems, especially simplex-based cutting plane and Lagrangian relaxation solution methods has been presented.

Adding capacity to the arcs of network design problem adds more complexity to this problem [2]. Furthermore, large-scale network design problems, which occur in real-world applications, are very difficult to solve [38]. There are theoretical and empirical evidence that the capacitated network design problems are NP-hard [38] [2]. It means that there is no efficient algorithm, which can solve them in polynomial time.

As a result, the researchers proposed approximate methods to solve them. However, this type of solution methods cannot guarantee the optimality of solutions. Heuristics and metaheuristics are the two classes of approximate methods [50]. Recent research works on developing approximate methods for multi-commodity network design problem, can be found in [9], [10], [35], and [39]. Heuristic solution method more often stop at the local optima. Hence, metaheuristics deal with these problems by introducing systematic rules to escape from local optima [14] [12] [44] [29].

The popular metaheuristics for MCND problems include Genetic Algorithm (GA) [32] [30] [15], Ant Colony Optimization (ACO) [13] [14], Tabu Search (TS) [27] [28], and Simulated Annealing (SA) [36]. Over the past two decades, Genetic algorithms have become as one of the solution method for some variants of network design problem [21], [22] [23] [24] [25] [37]. Since ACO concepts are suitable for network problems, one can find applications of that for MCND problems [48] [45]. In the case of TS algorithm, there are also several research works on MCND problems including [7], [26] [53], [46], [31], and [5]. The SA algorithm is another metaheuristic that is used for solving the MCND. In [33] a SA solution method for distribution network problem was proposed. In [48] a hybrid SA and ACO approach for solving transportation network design was proposed. In [47] a simulated annealing method for reverse logistic network design problem was proposed. In [34] a SA method for logistic network design problem was proposed. Other applications of SA in computer network design can be found in [52] and [16].

This paper proposes a SA algorithm as a metaheuristic solution method for unsplittable MCND problem. In addition, an innovative representation is proposed for representing solutions. Several instances are selected from OR-Library [3] for performance evaluation of proposed algorithms. The parameter tuning for the proposed algorithm is done by means of Design of Experiments (DOE) approach and Design-Expert statistical software [51]. The results of the algorithm are compared with CPLEX solver with branch-and-bound method.

The paper is organized as follows. In the next section, unsplittable MCNDP optimization model is presented. In section 3, the SA solution method to solve unsplittable MCNDP and its computational results are described. Finally, in section 4, the conclusion is presented.

2. Unsplittable MCND Problem

In unsplittable MCND problem, each commodity must follow exactly one rout between its origin and destination. It means, the flow variable must be a binary or 0-1 variable. Suppose a directed network $G = (N, A)$, in which $N$ is the set of nodes and $A$ is the set of directed arcs. Let $K$ be the set of commodities, then the amount of each commodity $k$ which must flow from its origin $O(k)$ to its destination $D(k)$ is $d_k$. Let $c_{ij}$ and $f_{ij}$ be the per unit arc routing cost of commodity $k$ on arc $(i, j)$, and fixed arc designing cost of arc $(i, j)$, respectively. The unsplittable MCND problem can be formulated as follows:

$$\min \sum_{k \in K} \sum_{(i, j) \in A} c_{ij} x_{ij}^k + \sum_{(i, j) \in A} f_{ij} y_{ij}$$

subject to

$$\sum_{j \in S} x_{ij}^k - \sum_{j \in S} x_{ij}^k = \begin{cases} 1 & \text{if } i = O(k) \\ -1 & \text{if } i = D(k) \text{ for all } K \in k \\ 0 & \text{otherwise} \end{cases}$$

$$\sum_{j \in S} x_{ij}^k d_{ij} \leq u_{ij} y_{ij} \quad \text{for all } (i, j) \in A$$

$$(x, y) \in S$$
\[ x^j_i, \quad y^j_i = 0 \text{ or } 1 \quad \text{for all} \quad (i, j) \in A, \quad k \in K \quad (5) \]

Where, \( y^j_i \) and \( x^j_i \) are decision variables. \( y^j_i \) is 0 if arc \((i, j)\) is closed, and 1 if it is open. \( x^j_i \) is 1 if commodity \( k \) is shipped on arc \((i, j)\), otherwise is 0. The objective function (1), is to minimize variable and fixed costs. Constraints (2) are the usual balancing equations of network flow problem, which insure that each demand is starting from origin node, passing intermediate nodes, and reaching to destination node. Constraints (3) demonstrate that the sum of flows on each arc \((i, j)\) must not exceed the capacity \( u^j_i \) of the arc. \( S \) is the set of side constraint, which can be added to this general model.

These constraints might be for example the topological restrictions or multiple-choice restrictions of the arcs. Unsplittable MCND, which is presented above is arc-based formulation.

3. The Proposed SA for the Unsplittable MCND

The SA was proposed by [36]. The overall procedure of SA is as follows. The SA start with initial solution and move iteratively to neighborhood solutions, if the neighborhood solution is better than current solution, it set as a current solution, otherwise it accepted by the probability \( \exp(-\Delta E/T) \) in which \( \Delta E \) is the difference between current and neighborhood solutions, and \( T \) is a parameter, which called temperature. At each temperature, several iterations are executed and then temperature is gradually decreased. At the first step of the search, temperature is set to high value (more probability of accepting worst solutions), and gradually it is decreased to a little value at the end of the search (less probability of accepting worst solutions).

3.1. The Solution Representation

The SA needs a representation of solutions, which constructs a solution and the neighborhood structure is defined based on it. For the proposed SA method for unsplittable MCND problem, we inspired representation from [22] in which a representation for genetic algorithm for shortest path problem was proposed. To illustrate proposed neighborhood structure, suppose figure 1 as a partial network of unsplittable MCND problem with 3 commodities. The capacity of each arc has been shown on it. Table 1 shows origin, destination and demand of each commodity \( k \). We use a \( m \times n \) matrix to represent a solution, in which \( m \) and \( n \) is related to commodities and nodes, respectively. Each value \((i, j)\) of this matrix shows a priority of node \( j \) for commodity \( i \). For example, in figure 2a, the value of element \((1, 2)\) is 5, that shows the priority of node 2 is 5 for commodity 1.

To obtain a solution from this structure, we must decode each matrix value. To do this, we start from origin node in each commodity, and check its unvisited neighborhood, the neighborhood that has maximum priority must be the next node in the path. For example, in figure 2, for commodity 1, the priority of neighborhoods of node 1, i.e., nodes 2, 3, and 4 are 5, 3, and 9, respectively, so that the next node must be node 4.

This must be done for each next node in each commodity to reach destination node. By this procedure, the resulting solution has been shown in figure 2b.

The capacities and flows have been shown on each arc. This proposed representation, in some cases, may lead to paths that are not reached to their destinations. To handle this problem, in proposed algorithm, solution is checked and if this problem occurs, the related solution is eliminated and a new solution will be constructed.

![Fig. 1. A network example of MCNDP](image)

| Tab. 1. Origin, Destination and demand of each commodity \( k \) |
|-----------------|---|---|
| \( O(k) \)     | \( D(k) \) | \( d_k \) |
| \( k_1 \)       | 1  | 13 | 2  |
| \( k_2 \)       | 2  | 11 | 2  |
| \( k_3 \)       | 4  | 12 | 2  |

3.2. Neighborhood Structure

To move from a current solution to a neighborhood solution, we use swap in a fraction of commodities. To do this, a number of commodities are selected randomly for swap. This is a user-defined number, which can be seen as a parameter of the algorithm. In each selected commodity, two nodes are selected randomly for swap. A swap is done by changing the positions of priorities. This process has been shown in figure 3 in which commodity 1 is selected for swap. In this commodity,
nodes 2 and 11 are selected randomly to swap. Resulting path of each solution is shown below it. In the special cases, these swaps might not lead to new solutions; i.e. there might be a situation that these swaps do not change a solution. The algorithm results show that in large networks with numerous arcs, this problem rarely occurs [21] and by selecting a fraction of commodities for swap instead of a single commodity there will be the low probability of occurrence this problem.

![Diagram](image_url)

Fig. 2. (a) Representation of a solution (b) Resulting paths for commodities

![Diagram](image_url)

Fig. 3. Neighborhood structure for SA for network example of figure 1

3.3. Evaluating Objective Function and Handling Constraints

For evaluating objective function, the flows of each arc is calculated according to solution and fixed costs are added for opening arcs in addition to variable costs according to the flows. The main challenges of MCND problems are how to handle constraints. For unsplittable MCND problem, the sum of flows on each arc must not exceed the maximum capacity of them. To handle this constraint, after generating a solution, the sum of flows on each arc must be checked. If this value is more than maximum capacity, then a penalty value is added to objective value of the solution, so that the probabilities of acceptance these solutions are decreased. This penalty is added according to below equation:

\[
(u_{y_i} - \sum_{i,j} x_{ij}^d)w_j, \ for \ (i,j) \in A \sum_{i,j} x_{ij}^d > u_{y_i}, \quad (6)
\]

where \(w_j\) is a user defined weight that represent penalties. This value can be seen as a parameter of the algorithm that must be tuned. This proposed penalty structure changes penalties according to solution. The worst the solution is, the more penalty is added. For illustration, consider figure 2 in which the flow of arcs (6-10) and (10-12) exceed the maximum capacity of the arcs. Therefore, we must add a penalty to objective function of this solution.
3.4. The Overall Structure of the Proposed SA

Figure 4 shows a pseudocode for the proposed SA method for unsplittable MCND problem.

```
Input
x = x0 ; /*create initial solution*/
T = T_{max} ; /*set initial temperature*/
Repeat
  Repeat /*in a constant temperature*/
  Swap (x = x') ; /*Generate a random neighbor*/
  Calculate f(x') and f(x);
  \Delta E = f(x') - f(x)
  If \Delta E <= 0 then x = x' /*accept neighborhood solution*/
  Else accept x' with a probability exp(-\Delta E / T) ;
  Until Equilibrium condition
  T = g(T) ; /*update temperature*/
  Until Stopping criteria satisfied /*e.g. T < T_{min}*/
Output: Best solution found.
```

Fig. 4. Proposed SA pseudo code

Creating initial solution is done randomly. In swap procedure, current solution is moved to neighborhood solution according to swap structure we mentioned above (figure 2). To calculate f(x') and f(x), the related matrix must be decoded into solutions and then the objective function must be calculated. If the solution is infeasible according to capacity constraint, a penalty must be added to the objective function. Initial temperature, annealing function for updating temperature, and the other parameters are tuned in the next section.

3.5. Parameters Tuning

The parameters of the proposed algorithms are tuned using the Design of Experiments (DOE) approach and Design-Expert statistical software [51]. One can define an experiment as a test or series of tests in which purposeful changes are made to the input variables of a process or system so that we may observe and identify the reasons for changes that may be observed in the output response.

DOE refers to the process of planning the experiment so that appropriate data that can be analyzed by statistical methods will be collected, resulting in valid and objective conclusions [4].

The three basic principles of DOE are replication, randomization, and blocking. Replication means a repetition of the basic experiment. Replication has two important properties. First, it allows the experimenter to obtain an estimate of the experimental error. Second, if the sample mean is used to estimate the effect of a factor in the experiment, replication permits the experimenter to obtain a more precise estimate of this effect.

Randomization means that both the allocation of the experimental material and the order in which the individual runs or trials of the experiment are determined randomly. Randomization usually makes this assumption valid. Blocking is a design technique used to improve the precision with which comparisons among the factors of interest are made.

Often blocking is used to reduce or eliminate the variability transmitted from nuisance factors; that is, factors that may influence the experimental response but in which we are not directly interested [43]. The important parameters in DOE approach are response variable, factor, level, treatment and effect. The response variable is the measured variable of interest. In the analysis of metaheuristics, the typically measures are the solution quality and computation time [1]. A factor is an independent variable manipulated in an experiment because it is thought to affect one or more of the response variables.

The various values at which the factor is set are known as its levels. In metaheuristic performance analysis, the factors include both the metaheuristic tuning parameters and the most important problem characteristics [4]. A treatment is a specific combination of factor levels. The particular treatments will depend on the particular experiment design and on the ranges over which factors are varied. An effect is a change in the response variable due to a change in one or more factors [49].

Design of experiments is a tool that can be used to determine important parameters and interactions between them. Four-stages of DOE consist of screening and diagnosis of important factors, modeling, optimization and assessment. This methodology is called sequential experimentation, which is used to set the parameters in the DOE approach and has been used in this paper for the proposed algorithms [43].

Several samples with different characteristics are selected from OR-Library for parameter tuning and performance evaluation of this algorithm. These instances are randomly generated benchmark instances described in [18] and [19]. The same instances have been used in several other papers, such as [7], [10], [26], [11], [35], [46], and [39].

The proposed algorithm is implemented with Java programming language. Table 2 shows these samples and their characteristics. In the table, problems are sorted according to the number of commodities. In the two last columns, F means that the fixed costs are predominate relatively to the variables costs and V means the converse, while L means that the capacities are loose while T means that the capacities are tight.

International Journal of Industrial Engineering & Production Research, June 2012, Vol. 23, No. 2
are examined for this function to reach the best value of it.

\[
\text{relative gap} = \frac{\text{obtained solution} - \text{optimum(lower bound)}}{\text{optimum(lower bound)}} \quad (7)
\]

We use geometric cooling function for decreasing temperature:

\[
T = \alpha T, \quad (8)
\]

where \( T \) and \( \alpha \) are temperature and cooling rate, respectively. As illustrated in table 3, several \( \alpha \) values

<table>
<thead>
<tr>
<th>Problem name</th>
<th>Number of nodes</th>
<th>Number of arcs</th>
<th>Number of Commodity</th>
<th>F/V</th>
<th>L/T</th>
</tr>
</thead>
<tbody>
<tr>
<td>C35</td>
<td>20</td>
<td>230</td>
<td>40</td>
<td>V</td>
<td>T</td>
</tr>
<tr>
<td>C36</td>
<td>20</td>
<td>230</td>
<td>40</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>C41</td>
<td>20</td>
<td>300</td>
<td>40</td>
<td>V</td>
<td>L</td>
</tr>
<tr>
<td>C43</td>
<td>20</td>
<td>300</td>
<td>40</td>
<td>V</td>
<td>T</td>
</tr>
<tr>
<td>C44</td>
<td>20</td>
<td>300</td>
<td>40</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>C49</td>
<td>30</td>
<td>520</td>
<td>100</td>
<td>V</td>
<td>L</td>
</tr>
<tr>
<td>C50</td>
<td>30</td>
<td>520</td>
<td>100</td>
<td>F</td>
<td>L</td>
</tr>
<tr>
<td>C57</td>
<td>30</td>
<td>700</td>
<td>100</td>
<td>V</td>
<td>L</td>
</tr>
<tr>
<td>C58</td>
<td>30</td>
<td>700</td>
<td>100</td>
<td>F</td>
<td>L</td>
</tr>
<tr>
<td>C37</td>
<td>20</td>
<td>230</td>
<td>200</td>
<td>V</td>
<td>L</td>
</tr>
<tr>
<td>C38</td>
<td>20</td>
<td>230</td>
<td>200</td>
<td>F</td>
<td>L</td>
</tr>
<tr>
<td>C39</td>
<td>20</td>
<td>230</td>
<td>200</td>
<td>V</td>
<td>T</td>
</tr>
<tr>
<td>C40</td>
<td>20</td>
<td>230</td>
<td>200</td>
<td>F</td>
<td>T</td>
</tr>
<tr>
<td>C45</td>
<td>20</td>
<td>300</td>
<td>200</td>
<td>V</td>
<td>L</td>
</tr>
<tr>
<td>C46</td>
<td>20</td>
<td>300</td>
<td>200</td>
<td>F</td>
<td>L</td>
</tr>
<tr>
<td>C47</td>
<td>20</td>
<td>300</td>
<td>200</td>
<td>V</td>
<td>T</td>
</tr>
<tr>
<td>C53</td>
<td>30</td>
<td>520</td>
<td>400</td>
<td>V</td>
<td>L</td>
</tr>
<tr>
<td>C54</td>
<td>30</td>
<td>520</td>
<td>400</td>
<td>F</td>
<td>L</td>
</tr>
<tr>
<td>C64</td>
<td>30</td>
<td>700</td>
<td>400</td>
<td>F</td>
<td>T</td>
</tr>
</tbody>
</table>

To tune parameters for the proposed algorithm we classified instances into 3 classes, according to the number of commodities. The instances that have 40, 100 to 200, and 400 commodities are categorized in classes 1, 2, and 3, respectively. For the classes 1 to 3, instances C35, C58, and C64 are chosen as representatives for parameter tuning. Parameters tuning is done separately for each group.

In the proposed algorithm, solution quality and CPU time are considered as the response variables. Factors, their levels, and the final obtained values are shown in Table 3. Each block is considered with 16 treatments and main effects. As a solution quality indicator, for each instance the average deviation from optimal value or lower bound result is calculated. Lower bound is used when CPLEX con not reach to optimal value.

3.6. Experimental Results

As we mentioned in pervious section, three of samples are used for parameter tuning. The algorithm is implemented with tuned parameter, for 16 remaining samples. The results are summarized in table 4. In column “CPLEX time”, “10h” means the CPLEX solver cannot reach to global optimum within 10 hours and CPLEX is stopped at this point.

Figure 5 compares computational time of CPLEX and the proposed algorithm. This shows a significant difference in computational time between CPLEX and the proposed algorithm. Figure 6 shows the relative gap in different samples.

This figure shows the gaps are increased as the size of samples increased. Figure 7 compares the objectives of CPLEX and SA. Figure 5 and 7 illustrate that SA can find near optimal solution in much less time.

International Journal of Industrial Engineering & Production Research, June 2012, Vol. 23, No. 2
Tab. 4. Results of implementing algorithm for samples

<table>
<thead>
<tr>
<th>Problem</th>
<th>CPLEX objective</th>
<th>CPLEX time</th>
<th>SA objective</th>
<th>SA time</th>
<th>Relative gap %</th>
</tr>
</thead>
<tbody>
<tr>
<td>C36</td>
<td>668699</td>
<td>0.41 sec</td>
<td>670928</td>
<td>9 sec</td>
<td>0.33</td>
</tr>
<tr>
<td>C41</td>
<td>429335</td>
<td>0.4 sec</td>
<td>433929</td>
<td>10 sec</td>
<td>1.07</td>
</tr>
<tr>
<td>C43</td>
<td>501766</td>
<td>0.4 sec</td>
<td>511384</td>
<td>8 sec</td>
<td>1.92</td>
</tr>
<tr>
<td>C44</td>
<td>643395</td>
<td>0.55 sec</td>
<td>656412</td>
<td>9 sec</td>
<td>2.02</td>
</tr>
<tr>
<td>C49</td>
<td>54387</td>
<td>7 min 23 sec</td>
<td>57422</td>
<td>1 min 24 sec</td>
<td>5.58</td>
</tr>
<tr>
<td>C57</td>
<td>47883</td>
<td>9 min 10 sec</td>
<td>51505</td>
<td>3 min 57 sec</td>
<td>7.56</td>
</tr>
<tr>
<td>C58</td>
<td>61204</td>
<td>10 h</td>
<td>68179</td>
<td>8 min 15 sec</td>
<td>11.40</td>
</tr>
<tr>
<td>C37</td>
<td>94111</td>
<td>10 h</td>
<td>118785</td>
<td>9 min 16 sec</td>
<td>26.22</td>
</tr>
<tr>
<td>C39</td>
<td>99099</td>
<td>10 h</td>
<td>112792</td>
<td>13 min 24 sec</td>
<td>13.82</td>
</tr>
<tr>
<td>C40</td>
<td>138355</td>
<td>10 h</td>
<td>167006</td>
<td>8 min 54 sec</td>
<td>20.71</td>
</tr>
<tr>
<td>C45</td>
<td>76441</td>
<td>10 h</td>
<td>91702</td>
<td>8 min 15 sec</td>
<td>19.96</td>
</tr>
<tr>
<td>C46</td>
<td>119047</td>
<td>10 h</td>
<td>151513</td>
<td>8 min 40 sec</td>
<td>27.27</td>
</tr>
<tr>
<td>C47</td>
<td>77006</td>
<td>10 h</td>
<td>88550</td>
<td>8 min 28 sec</td>
<td>14.99</td>
</tr>
<tr>
<td>C53</td>
<td>115414</td>
<td>10 h</td>
<td>134587</td>
<td>21 min 34 sec</td>
<td>16.61</td>
</tr>
<tr>
<td>C64</td>
<td>138712</td>
<td>10 h</td>
<td>169070</td>
<td>20 min 14 sec</td>
<td>21.16</td>
</tr>
</tbody>
</table>

Fig. 5. Computational time for CPLEX and SA in different samples

Fig. 6. Relative gap of SA in different samples
4. Conclusions
In this paper, we have proposed a SA solution method with an innovative solution representation and neighborhood structure for unsplittable multicommodity capacitated network design problems. The results show that SA can find near optimal solution in much less time than exact algorithm. We hope proposed SA approach for unsplittable MCND problem will draw more researchers to develop and implement this approach for real-life problems. In addition, the proposed representation and neighborhood structure can be used in another metaheuristics such as genetic algorithm, and tabu search, so that, it can be another new search field.

References


A Simulated Annealing Algorithm for Unsplittable


