Mathematical and Computer Modelling 57 (2013) 382–400

Contents lists available at SciVerse ScienceDirect

Mathematical and Computer Modelling

journal homepage: www.elsevier.com/locate/mcm

Modeling the budget-constrained dynamic uncapacitated facility location–network design problem and solving it via two efficient heuristics: A case study of health care

Abdolsalam Ghaderi *, Mohammad Saeed Jabalameli

Department of Industrial Engineering, Iran University of Science and Technology, P.C. 1684613114, Narmak, Tehran, Iran

ARTICLE INFO

Article history:
Received 21 February 2011
Received in revised form 12 December 2011
Accepted 19 June 2012

Keywords:
Facility location
Network design
Dynamic (multi-period)
Health care
Fix-and-optimize heuristic

ABSTRACT

This paper presents a model for the budget-constrained dynamic (multi-period) uncapacitated facility location–network design problem (DUFLNDP). This problem is concerned with the determination of the optimal locations of facilities and the design of the underlying network simultaneously in which there is a budget constraint on investment for opening the facilities and constructing (activating) links for each time period during the planning horizon. The objective is to minimize the total travel costs for customers and operating costs for facilities and network links. Furthermore, a greedy heuristic and a fix-and-optimize heuristic based on simulated annealing and exact methods (Branch & Bound and cutting methods) are proposed to solve the model. The performance of the proposed algorithms were tested on extensive randomly generated instances and compared with the CPLEX solver. The budget-constrained DU FLNDP has a lot of important applications which as one of them, the accessibility of health care facilities in Illam Province of Iran is investigated.

1. Introduction

Generally, facility location problems deal with the decisions of where to optimally locate facilities (factories, distribution centers, warehouses, schools, hospitals, etc.) and how to allocate customers to facilities such that the demand for some service or product is satisfied. In addition, network design problems deal with the decisions about network links construction and determining traffic flow on these links and possibly satisfy additional constraints. Usually, these decisions are made by considering the associated costs (or profits) of satisfying the demand and the costs related to establishing (or operating) the facilities and links. Essentially, the facility location–network design problem is a combination of the facility location and network design that involves the determination of the location of the facilities (as in facility location) required to satisfy a set of clients' demands and the determination of travelable links (as in network design) to connect clients to facilities.

Unlike most classical location models, two main assumptions are considered at the facility location–network design problem. First, the network is not given and the model determines the configuration of the underlying network. In the classical models, facilities are located on a given network. In addition, each client is not directly served from facilities and may pass from multiple nodes in the network to get service. However, this subject is not often established in the most classical location models.

Many efforts have been advocated on developing models and applications in each of these two fields of science. The goal here is not to review all related works in the literature, but to mention key references to clarify the problem at hand.

* Corresponding author. Tel.: +98 9128085385.
E-mail addresses: ab_ghaderi@iust.ac.ir, ab.ghadri@gmail.com (A. Ghaderi).

0895-7177/$ – see front matter
doi:10.1016/j.mcm.2012.06.017

Weber [1] was the first one to formulate the facility location in 1909. After that, many papers have been published that provide admirable introductions and reviews of the development in the field. One may refer to survey papers in this area such as [2–6]. In addition, the network design problem is relatively newer than the facility location problem. Magnanti and Wong [7] and Yang and Bell [8] proposed a general problem description of the network design problem that encompasses many variations. Furthermore, they presented an outstanding review of the network design problem models and discussed the solution methods.

However, the facility location–network design problem has been considered less in the literature. The uncapacitated facility location–network design problem (UFLNDP) was originally proposed by Daskin et al. [9]. Later, Melkote [10] in his doctoral thesis investigated three models for the facility location–network design problem including: UFLNDP, the capacitated facility location–network design problem (CFLNDP), and the maximum covering location–network design problem (MCLNDP). The results of this thesis were published in [11,12]. Drezner and Wesolowsky [13] proposed a new network design problem with potential links where each link can be either constructed or not at a given cost. In addition, each constructed link can be constructed either as a one-way or two-way link. In the paper, four basic problems were created subject to two objective functions. Then, these problems were solved by a descent algorithm, simulated annealing, tabu search, and a genetic algorithm. In another doctoral thesis, Cocking [14] developed many approaches to solve the static budget constrained facility location–network design (FLND) problem. A number of algorithms were introduced to find good upper bounds and good lower bounds on the optimal solution value. The developed heuristics were: simple greedy heuristics, a local search heuristic, metaheuristics including simulated annealing (SA) and variable neighborhood search (VNS), as well as a custom heuristic based on the problem-specific structure of FLND. Furthermore, a branch-and-cut method that uses heuristic solutions as upper bounds, and cutting planes for increasing the lower bound at each node of the problem tree were developed. This approach reduces the number of nodes needed to solve to optimality. Some of the results from this thesis were published in [15].

Recently, Bigotte et al. [16] proposed a mixed-integer optimization model for integrated urban hierarchy and transportation network planning. This model simultaneously determines which urban centers and which network links should be promoted to a new level of hierarchy to maximize accessibility to all classes of facilities.

Based on some assumptions such as solution space, uncertainty in parameters, and time horizon, the facility location problems can be classified into different categories. Most facility location models available in the literature assume that demand and costs are known and do not change over time. Once the facilities have been optimally located, they are assumed to remain sited regardless of how demand and costs may change in the future time periods. These models are called single-period or static location models. In practice, however, the demand is unknown and is time varying. In addition, other parameters such as the transportation and operation costs may increase or decrease from one period to another. Thus, if the problem parameters are time varying, it might be necessary to relocate the facilities to meet the upcoming changes. The dynamic location models have been developed in the literature to overcome this subject. In these models, the optimal time and location of facilities are determined such that the total costs for serving demand and for operating and relocating the facilities are minimized.

Current et al. [17] has divided dynamic location models into explicit and implicit models. In explicit models, the opening and closing of facilities can occur during the planning horizon subject to changing at the problem parameters. In implicit models, all facilities are to be opened at the beginning of the planning horizon and remain open throughout. Nevertheless, these models consider changing in parameters in the initial set of locations. Manne [18,19] and Ballou [20] were researchers that proposed the first dynamic models of the facility location. After that, much work has been done on the dynamic facility location and many models have been developed. These models include but are not limited to the dynamic single facility location problem [20,21], dynamic multi-facility location problems [22,23], the dynamic p-median problem [24,25], set-covering problem [26,27], location-routing problem [28], and dynamic multi-echelon facility location problem [29,30]. Here, only some related papers in the dynamic area are mentioned and the reader is referred to [31,32] for a comprehensive review of this literature. Furthermore, the dynamic network design problem is also considered in the literature. For example, one may refer to [33,34] to read more information.

In this paper, a dynamic version of UFLNDP with time varying parameters is investigated. All of the previous works typically consider this problem in a static (or a single time period) state. The dynamic budget-constrained uncapacitated facility location–network design problem aims at finding facilities opening and links construction (improvement) during the planning horizon so as to minimize travel and operating cost subject to the budget constraint. In addition, a discrete and finite time horizon is assumed in this problem. There are fixed costs related to opening new facilities and network links. In addition, there are operating costs of the existing facilities and links and a transportation cost for serving the demand of customers from open facilities. Furthermore, the facilities and network links can be opened at different time periods to provide a flexible network.

In general, the investigated problem is useful to model a number of real-world applications in which tradeoffs between facility costs, network design costs, and operating costs must be made. Such applications arise in the regional planning problems, automated guided vehicles (AGVs) systems, pipeline systems, power transmission networks, airline networks, and telecommunications networks, to name just a few contexts. For example, regional planning perhaps is the most direct application of this model to design a new access infrastructure or to improve an existing network. In such a setting, the government may be simultaneously considering the construction of the new roadway system as well as the location of public and government facilities such as post office, schools, and fire stations with a limited budget [10]. Cocking [14] studied such a specific real-world setting with its goal to improve access to health facilities for the people in the Nouna health district of Burkina Faso. Remarkable improvements in accessibility to health facilities were obtained.

Clearly, this type of problem needs a lot of expenditure for investment. Thus, in some practical situations, the decisions about the location of facilities and links may not be possible due to the budget constraint. Such situations may arise in the public or private sectors in which facilities like power plants, hospitals, and schools are expected to be operating for a long period of time.

Both the facility location and network design problems as subproblems of the facility location–network design problem are NP-hard. Hence, the budget constrained DUFLNDP is clearly NP-hard since it combines two NP-hard problems. Many algorithms and solution approaches have been developed to solve the dynamic facility location and network design problems. In this paper, a heuristic and a fix-and-optimize heuristic are proposed to solve the problem under study. The fix-and-optimize approach is an efficient hybrid algorithm that combines an exact optimization method with simulated annealing algorithm based on the neighborhood structure. The proposed Fix-and-optimize algorithm differs somewhat from one was proposed by Pochet and Wolsey [35] and Helber and Sahling [36] and Sahling et al. [37]. This is an iterative hybrid method that fixes some variables at each iteration and finds the problem solution optimally by an exact optimization method. Then, the fixed variables are searched by the SA algorithm.

The studied facility location–network design models in the literature ignore the time varying behavior of demand and cost parameters. The budget-constrained dynamic facility location–network design model in this paper addresses this limitation by incorporating strategic decisions to locate facilities and establishing network links throughout the time horizon. At a glance, this study has some contribution to the literature in the facility location–network design as follows.

- The proposed mathematical model for the budget-constrained DUFLNDP is novel in considering the opening of facilities and the network links with fixed costs and operating of facilities and links, simultaneously. Despite many published works on the dynamic facility location and network design separately in the literature, the budget-constrained DUFLNDP has not been investigated before.
- An approach is illustrated to formulate the budget constraint in the model such that the available budget at each time period is considered separately for the facility location and network design. This situation may happen in some real-world applications when two different organizations are responsible for investment in facilities and network design. In addition, the unspent budget at the end of time periods is not returned and may be used in the subsequent time periods.
- Two efficient heuristics are proposed to solve the model. At the first heuristic, the problem is divided into two much easier sub-problems and each of them is solved by CPLEX. In addition, the second heuristic is an efficient hybrid algorithm based on simulated annealing. The performance of the proposed algorithms is compared with CPLEX solver for various test problems in different scales.
- An application of the budget-constrained dynamic uncapacitated facility location—network design is presented to improve access to health facilities for the rural population centers in the Illam Province of Iran.

The remaining sections of this paper are organized as follows. In Section 2, the mathematical formulation of the proposed problem is presented. A brief description of the heuristic and hybrid SA algorithm is given in Section 3. In Section 4, the model is applied to a case study in the Illam Province of Iran and some other computational studies with generated data are reported to demonstrate the efficiency of proposed algorithms. Finally, conclusions and directions for further research are given in Section 5.

2. Problem statement and mathematical formulation

2.1. Problem definition and assumptions

The budget constrained DUFLNDP is defined on a network made by a set of clients, a set of potential locations for facilities, a set of potential links for constructing the network, and a time horizon described in terms of a set of consecutive time periods. In addition to some assumptions that were considered for the UFLNDP by Daskin et al. [9], the following assumptions are also considered: (1) the facilities and links are uncapacitated, (2) parameters change over time with a specific process, (3) in most applications including health care, it is not likely to build a link or a new hospital in one year, close it at the end of that year, and then rebuild it two years later. Thus, it is assumed that once a facility or a link is opened, it remains open throughout the time horizon and will never be closed until the end of the planning horizon, (4) opening of facilities as well as constructing of links are instantaneous, (5) opening of a link or facility must happen at the beginning of a time period.

2.2. Notations

The model of the budget constrained DUFLNDP is formulated below as a mixed integer non-linear programming (MINLP) problem. To simplify the mathematical formulation of the problem, the following notations shown in Table 1 are considered to model the proposed problem.

The proposed model can be implemented in two scenarios: (1) Construct a new network and (2) Improve an existing network. In these problems, the impact of having control over both facility location and network design decisions during

Table 1
Notations used in the proposed model.

Symbol	Description
Sets	
N	Set of network nodes, $i, j \in \{1, 2, \ldots, \|N\|\}$ and set of clients, $k \in \{1, 2, \ldots, \|N\|\}$,
N^0	Set of opened facilities in existing network, $N^0 \in \{1, 2, \ldots, \|N\|\}$,
LE^t	Set of existing links at time period t, $(i, j) \in LE^t$,
LP^t	Set of potential links at time period t, $(i, j) \in LP^t$,
L^t	Set of network links at time period t, $(i, j) \in L^t$, $L^t = LE^t \bigcup LP^t$,
L^0	Set of constructed links in existing network, $(i, j) \in L^0$,
T	Set of time periods, $t \in \{1, 2, \ldots, \|T\|\}$,
Parameters	
d_k^t	The demand of client k at time period t,
m_{ij}	The length of link (i, j),
g_i^t	Fixed cost of opening a facility on node i at time period t,
c_{ij}^t	Fixed cost of constructing a link (i, j) at time period t,
ρ_{ij}^t	Travel cost per unit flow on link (i, j) at time period t,
r_{ij}^{kt}	The cost of traveling on link (i, j) if all the demand of client k goes through that link at time period t, $r_{ij}^{kt} = \rho_{ij}^t d_k^t$,
f_i^t	Operating cost of opened facility on node i during time period t,
h_{ij}^t	Operating cost of constructed link on (i, j) during time period t,
$\bar{B}^t$	Available budget for investing in facilities at time period t,
$\hat{B}^t$	Available budget for investing in network links at time period t,
Decision variables	
Z_i^t	If facility i is open at the beginning of time period $t(1)$, otherwise (0),
X_{ij}^t	If link (i, j) is open at the beginning of time period $t(1)$ otherwise (0),
Y_{ij}^{kt}	Fraction of the client's demand k traveling i to j at time period t,
W_i^{kt}	Fraction of the client's demand k served by facility i at time period t.

the planning horizon is measured. In the model, time periods start with 1, and $t = 0$ represents what is already in existence. N^0 is the set of facilities already opened prior to the first time period, or one can think of N^0 as the facilities opened at $t = 0$. Likewise, L^0 is the set of links already opened prior to the first time period or open at time 0.

The decisions of the dynamic facility location and network design configuration include deciding the facilities and links that are to be built at the potential locations and arcs as well as the quantities of clients' demand that are to travel on the transportation links in each time period.

2.3. Model formulation

Using these notations and assumptions, the mathematical formulation of the budget-constrained DUFLNDP is shown below:

$$\textbf{Min} \sum_{t\in T} \sum_{(i,j)\in L^t} \sum_{k\in N} r_{ij}^{kt} Y_{ij}^{kt} + \sum_{t\in T} \sum_{i\in N} f_i^t Z_i^t + \sum_{t\in T} \sum_{(i,j)\in L^t : i<j} h_{ij}^t X_{ij}^t. \tag{1}$$

Subject to

$$Z_i^t + \sum_{j\in N} Y_{ij}^{it} = 1 \quad \forall i \in N,\ \forall t \in T, \tag{2}$$

$$\sum_{j\in N} Y_{ji}^{kt} = \sum_{j\in N} Y_{ij}^{kt} + W_i^{kt} \quad \forall i, k \in N : i \neq k,\ \forall t \in T, \tag{3}$$

$$Z_k^t + \sum_{i\in N: i\neq k} W_i^{kt} = 1 \quad \forall k \in N,\ \forall t \in T, \tag{4}$$

$$Y_{ij}^{kt} + Y_{ji}^{kt} \leq X_{ij}^t \quad \forall (i, j) \in L^t : i < j,\ \forall k \in N,\ \forall t \in T, \tag{5}$$

$$W_i^{kt} \leq Z_i^t \quad \forall i, k \in N : i \neq k,\ \forall t \in T, \tag{6}$$

$$\sum_{t'=1}^{t} \sum_{i\in N} g_i^{t'} Z_i^{t'} (1 - Z_i^{t'-1}) \leq \sum_{t'=1}^{t} \bar{\beta}^{t'} \quad \forall t \in T, \tag{7}$$

$$\sum_{t'=1}^{t} \sum_{(i,j)\in L^t: i<j} c_{ij}^{t'} X_{ij}^{t'} (1 - X_{ij}^{t'-1}) \le \sum_{t'=1}^{t} \hat{\beta}^{t'} \quad \forall t \in T, \tag{8}$$

$$Z_i^t \ge Z_i^{t-1} \quad \forall i \in N,\ \forall t \in T, \tag{9}$$

$$X_{ij}^t \ge X_{ij}^{t-1} \quad \forall (i,j) \in L^t : i < j, \forall t \in T, \tag{10}$$

$$Z_i^{t-1} = 1 \quad \forall i \in N^0 \quad \text{and} \quad Z_i^{t-1} = 0 \quad \forall i \notin N^0,\ \forall t = 1, \tag{11}$$

$$X_{ij}^{t-1} = 1 \quad \forall (i,j) \in L^0 : i < j \quad \text{and} \quad X_{ij}^{t-1} = 0 \quad \forall (i,j) \notin L^0 : i < j,\ \forall t = 1, \tag{12}$$

$$Y_{ij}^{kt} \ge 0 \quad \forall (i,j) \in L^t,\ \forall k \in N,\ \forall t \in T, \tag{13}$$

$$X_{ij}^t \in \{0, 1\} \quad \forall (i,j) \in L^t : i < j,\ \forall t \in T, \tag{14}$$

$$W_i^{kt} \ge 0 \quad \forall i, k \in N : k \ne i,\ \forall t \in T, \tag{15}$$

$$Z_i^t \in \{0, 1\} \quad \forall i \in N,\ \forall t \in T. \tag{16}$$

The objective function (1) includes the total cost over the time horizon and minimizes only the traveling and operating costs. These costs are composed of the two main components, the sum of transportation costs and the operating costs of the facilities and of the network. However, investment costs are not considered in the objective function. Investment costs, which include the opening cost of facilities at potential locations and the construction cost of links at potential links, should be subject to the maximum available budget in each time period.

Eqs. (2)–(4) are the flow conservation conditions, which must hold for each client, facility and period. Unlike Melkote and Daskin's model [11] where $Y_{kj}^k = X_{kj}$, this equation is not necessarily satisfied here because arc (k, j) may have been previously constructed (if $X_{kj}^{t-1} = 1$ then $X_{kj}^t = 1$) and there is no flow of client k on arc (k, j) in the current time period ($Y_{kj}^{kt} = 0$). Thus, Y_{kj}^{kt} is not necessarily equal to X_{kj}^t.

Constraint (2) ensures that demand at i is either served by a facility at i or by shipping on some link out of i. Once a link is constructed, it remains open throughout the time horizon. Constraint (3) states in time period t and for client k that the flow into i must equal the flow out of i. Constraint (4) imposes that in time period t and for client k, the demand must find a destination, whether it be at node k itself (Z_k^t) or at the other nodes i (W_i^{kt}).

Constraints (5) and (6) guarantee that in each time period, potential links and facilities are not used if they are not constructed. Constraint (5) is equivalent to ones in UFLNDP that says on any given link, an optimal solution flow for a given time period will be in only one direction. Therefore, the links (i, j) and (j, i) cannot both be constructed in the static case. However, in the dynamic case, once a link is constructed, it remains open throughout the time horizon. Therefore, at some time periods, the passing flow on some constructed links may be zero. This situation occurs when these links were built previously. In this model, we model the link opening variables as undirected, and ask only that if a link is used in either direction, then it is open.

Inequalities (7) and (8) are the budget constraints for investing in facilities and network that are described in detail in the following sub-section. Constraints (9) and (10), respectively imply that once a facility and link is constructed, it remains open throughout the time horizon. Constraints (11) and (12) establish "$t - 1$" for the first time period, i.e., $t = 1$ is the first time period, respectively, that exists prior to that at $t = 0$. Constraints (13) and (15) enforce the non-negativity of the flow variables. Constraints (14) and (16) enforce the binary restriction on the location and link decision variables.

According to the single assignment property, every client's demand is completely assigned to a single facility (the closest). That is, nothing is gained by "splitting up" a demand and sending parts of it to different facilities. Hence, as long as the demands are integer-valued, Y_{ij}^{kt} and W_i^{kt} are integral [11].

2.3.1. Budget constraints

In the real-world applications, subject to the organizations involved in investment at the facility location and network design and their policies, some situations may arise where the available budget needs to be considered. One of these policies may be whether the unspent budget at the end of time periods is returned or not. In addition, a single or more organizations may be responsible for investing in the facility location and network design. Hence, the combination of these policies could be employed to model the budget constraint in the studied problem. In this paper, it is assumed that two different organizations are responsible for investment in facilities and network links. Thus, the total investment budget of facilities and network is considered separately. Moreover, the unspent budget at the end of time periods is not returned and may be used in the subsequent time periods. The related budget constraints of this approach are obtained as follows and represented as (7) and (8). These constraints state that the total investment expenses in facilities and network until time period t should not exceed the cumulative available budget in that time period.

$$\forall\, t = 1: \quad \sum_{i\in N} g_i^1 Z_i^1 (1 - Z_i^0) \le \bar{\beta}^1$$

$$\forall\, t = 2: \quad \sum_{i\in N} g_i^2 Z_i^2 (1 - Z_i^1) \le \bar{\beta}^2 + \left(\bar{\beta}^1 - \sum_{i\in N} g_i^1 Z_i^1 (1 - Z_i^0)\right)$$

$$\rightarrow \quad \sum_{i\in N} g_i^1 Z_i^1 (1 - Z_i^0) + \sum_{i\in N} g_i^2 Z_i^2 (1 - Z_i^1) \le \bar{\beta}^2 + \bar{\beta}^1$$

$$\forall\, t = 3: \quad \sum_{i\in N} g_i^1 Z_i^1 (1 - Z_i^0) + \sum_{i\in N} g_i^2 Z_i^2 (1 - Z_i^1) + \sum_{i\in N} g_i^3 Z_i^3 (1 - Z_i^2) \le \bar{\beta}^1 + \bar{\beta}^2 + \bar{\beta}^3$$

$$\Rightarrow$$

$$\forall t \in T: \sum_{t'=1}^{t} \sum_{i\in N} g_i^{t'} Z_i^{t'} (1 - Z_i^{t'-1}) \le \sum_{t'=1}^{t} \bar{\beta}^{t'}.$$

And likewise for the links:

$$\forall\, t = 1: \quad \sum_{(i,j)\in L^t: i<j} c_{ij}^1 X_{ij}^1 (1 - X_{ij}^0) \le \hat{\beta}^1$$

$$\forall\, t = 2: \quad \sum_{(i,j)\in L^t: i<j} c_{ij}^2 X_{ij}^2 (1 - X_{ij}^1) \le \hat{\beta}^2 + \left(\hat{\beta}^1 - \sum_{(i,j)\in L^t: i<j} c_{ij}^1 X_{ij}^1 (1 - X_{ij}^0)\right)$$

$$\rightarrow \quad \sum_{(i,j)\in L^t: i<j} c_{ij}^1 X_{ij}^1 (1 - X_{ij}^0) + \sum_{(i,j)\in L^t: i<j} c_{ij}^2 X_{ij}^2 (1 - X_{ij}^1) \le \hat{\beta}^2 + \hat{\beta}^1$$

$$\forall\, t = 3: \sum_{(i,j)\in L^t: i<j} c_{ij}^1 X_{ij}^1 (1 - X_{ij}^0) + \sum_{(i,j)\in L^t: i<j} c_{ij}^2 X_{ij}^2 (1 - X_{ij}^1) + \sum_{(i,j)\in L^t: i<j} c_{ij}^3 X_{ij}^3 (1 - X_{ij}^2) \le \hat{\beta}^1 + \hat{\beta}^2 + \hat{\beta}^3$$

$$\Rightarrow$$

$$\forall t \in T\ : \quad \sum_{t'=1}^{t} \sum_{(i,j)\in L^t: i<j} c_{ij}^{t'} X_{ij}^{t'} (1 - X_{ij}^{t'-1}) \le \sum_{t'=1}^{t} \hat{\beta}^{t'}.$$

2.3.2. Model linearization

The studied budget-constrained DUFLNDP has a mixed-integer non-linear programming (MINLP) model regarding to some non-linear terms in the budget constraints. However, it can be easily linearized by introducing new binary variables and additional constraints as follows:

$$U_i^t = Z_i^t Z_i^{t-1}$$

$$V_{ij}^t = X_{ij}^t X_{ij}^{t-1}.$$

The following constraints:

$$U_i^t \le Z_i^t \quad \forall i \in N,\ \forall t \in T, \tag{17}$$

$$U_i^t \le Z_i^{t-1} \quad \forall i \in N,\ \forall t \in T, \tag{18}$$

$$U_i^t \ge Z_i^t + Z_i^{t-1} - 1 \quad \forall i \in N, \forall t \in T, \tag{19}$$

$$V_{ij}^t \le X_{ij}^t \quad \forall (i,j) \in L^t : i < j, \forall t \in T, \tag{20}$$

$$V_{ij}^t \le X_{ij}^{t-1} \quad \forall (i,j) \in L^t : i < j,\ \forall t \in T, \tag{21}$$

$$V_{ij}^t \ge X_{ij}^t + X_{ij}^{t-1} - 1 \quad \forall (i,j) \in L^t : i < j,\ \forall t \in T \tag{22}$$

as well as the equivalent budget constraints for the mixed integer linear programming (MILP) model are formulated in the following inequalities:

$$\sum_{t'=1}^{t} \sum_{i\in N} g_i^{t'} (Z_i^{t'} - U_i^{t'}) \le \sum_{t'=1}^{t} \bar{B}^{t'} \quad \forall t \in T, \tag{23}$$

$$\sum_{t'=1}^{t} \sum_{(i,j)\in L^t: i<j} g_{ij}^{t'} (X_{ij}^{t'} - V_{ij}^{t'}) \le \sum_{t'=1}^{t} \hat{B}^{t'} \quad \forall t \in T. \tag{24}$$

As a result, the final proposed model of the budget-constrained DUFLNDP is easily converted to MILP.

Step 1: Set $t=1$ and solve P^1,

Step 2: fix Z_i^1 and X_{ij}^1 to the optimal solution obtained from **Step 1** and solve ***P***

Fig. 1. Pseudo code of greedy heuristic.

3. Solution procedures

The proposed model was coded in Python and solved by CPLEX solver. Traditional solvers like CPLEX are not an efficient means to solve the proposed model, because the problem is a comprised of two NP-hard problems; thus, the performance of CPLEX is not viable to solve the problem. In this model, even finding a feasible solution for medium and large scale instances is challenging. Thus, two heuristic algorithms are applied to solve the model. In both heuristics, some variables are fixed and the relaxed problem is solved by CPLEX solver. The second heuristic is an iterative hybrid metaheuristic algorithm, which is directly based on the simulated annealing and the modified linear mixed-integer program. Ideas from the well-known SA methodology were used to avoid becoming stuck at a local optimum solution. In addition, an exact optimization algorithm that uses Branch & Bound and cutting plane methods in its framework was used to find the optimal solution of subproblems.

3.1. Greedy heuristic

In this heuristic, the problem is split into two sub-problems and each of these problems is solved by CPLEX solver. Computational results showed that solving the problem at hand is not very difficult if the time periods are restricted to just single period or some of the facilities and links are fixed. The greedy heuristic has two phases. At the first phase, the problem is solved just for the first time period without any other changes in the model and variables. Then, in the second phase, the problem is solved by fixing the Z_i^1 and X_{ij}^1 variables to the values obtained at the first phase. Assume that ***P*** is the original problem and P^1 is the original problem with $T = 1$. Fig. 1 shows the pseudo-code of the greedy heuristic.

3.2. The fix-and-optimize or hybrid SA heuristic

The reasoning behind the proposed hybrid algorithm for the budget-constrained DUFLNDP is that the number of binary variables determines the majority of the numerical effort in the Branch & Bound process required to solve the mixed-integer program. By contrast, the number of real-valued variables is of secondary importance. Hence, a series of subproblems that are derived from the budget-constrained DUFLNDP are solved iteratively and quickly in a systematic manner. For each subproblem, the binary facilities and links opening variables at the first time period are set to a fixed value such that the model constraints should not be violated. The way how these binary variables are iteratively determined is based on simulated annealing algorithm that will be described in detail in the following. In each iteration, the CPLEX solver is used to find the optimal solution of the subproblem.

Simulated annealing is a relatively old and effective metaheuristic aimed at solving combinatorial and global optimization problems. This algorithm is a local search that escapes from the local optima. The idea behind simulated annealing comes from the physical process of annealing, in which a solid is heated to a given temperature and then slowly cooled to achieve an optimal crystal structure. The SA algorithm first appears in the literature by Kirkpatrick et al. [38] and since then has developed both in its methods and its applications.

The fix-and-optimize proposed algorithm was developed from the basic SA principle. The overall flowchart of the hybrid heuristic is shown in Fig. 3 and is comprised of two main processes as fix (SA) and optimize (CPLEX solver) operators. Note that generating the initial solution and the neighborhood of the current solution is vital to have a good performance of SA. Many strategies were examined for each of these main items, and, finally, the best combination of them was chosen in the proposed algorithm.

3.2.1. Solution representation

The proposed model has four decision variable sets to determine the outputs of the problem. Only variables associated with the opening of the facilities at the first time period are used in the solution representation. Other variables are obtained by solving the sub-problems of the model with the exact optimization methods using the CPLEX solver. For instance, the representation of a solution in this paper is shown in Fig. 2. In general form, the represented solution has T time periods and N nodes. Thus, each solution is represented by a matrix with $N \cdot T$ arrays that the first N arrays are generated by SA algorithm and the other arrays are found through solving the model with CPLEX. Each cell will have a 1 indicating that the facility is open during the corresponding time period, and 0 if it is not opened during that time period.

3.2.2. Construction of an initial solution

The initial solution is usually generated randomly. However, this process may lead to low quality initial locations. This, in turn, affects the final solution of the algorithm. In this algorithm, the greedy heuristic is used to generate the initial solution that has better performance than the completely random case.

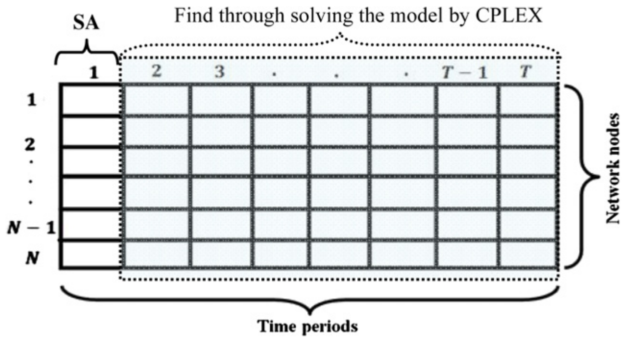

Fig. 2. A solution representation of Z_i^t variables for the proposed algorithm.

3.2.3. Neighborhood generating procedure

The performance of the proposed algorithms depends greatly on the neighborhood structure and other related operators. Thus, using an efficient procedure for neighborhood production is very important for a successful algorithm. In this paper, the neighborhood is obtained by swapping the value of two nodes that have different values at the first time period. For this purpose, the base moves for the problem are the interchange moves that swap a node in the solution with value $1(Z_i^1 = 1)$ by another one with value $0(Z_j^1 = 0)$. Hence, in the new solution, $(Z_j^1 = 1)$ and $(Z_i^1 = 0)$. In neighborhood generation, firstly, we look for a pair of nodes (i, j) with relationship $(i, j) \in L^t$ and then other pairs would be selected to swap their values. Then, with fixing the Z_i^1 variables for all nodes, the problem P^1 is solved by CPLEX to obtain the optimal network design variables. Finally, the problem is solved with fixing X_{ij}^1 and Z_i^1 variables and a neighborhood is generated with its objective function in this way.

3.2.4. Framework of the hybrid SA

The main idea behind this approach is that solving the problem at hand is not very difficult if the time periods restricted to just single periods or X_{ij}^1 and Z_i^1 variables are fixed. The main operations of the proposed algorithm are given in Fig. 3. At the first step of this scheme, the problem is solved by CPLEX with restricting the time periods to $T = 1$. By doing this, a feasible network with some links and facilities is constructed. Now, the value of X_{ij}^1 and Z_i^1 variables are fixed to the obtained value in the first time period and the problem is solved. Therefore, at the initialization step, the initial solution is set as the current solution. The current objective value is calculated by CPLEX and treated as the optimal objective value. The SA heuristic explores the feasible space and starts from the neighborhood of the current solution until a near-globally optimal solution is found or the stopping condition is met. Simulated annealing always accepts a better solution based on the objective function but it also reduces the likelihood of the solution being trapped, by accepting a worse solution if an acceptance criterion value is greater than a selected uniform random number.

Hence, the SA is applied to search on fixed variables that are Z_i^1. Then, the optimal solution of the sub-problems is obtained by CPLEX solver. Finally, the heuristic terminates and reports the best found solution. Parameters in the proposed framework of the fix-and-optimize heuristic are as follows:

- IT: initial temperature
- T: the current temperature
- NUM: the number of searched neighborhood at each temperature
- β: the cooling rate
- ΔE: the objective function difference between the current solution and its neighborhood
- r: a random value between [0, 1]
- K: a constant
- i: an counter.

As shown in Fig. 3, three sub-problems are solved with the exact algorithms by CPLEX solver in the proposed hybrid SA as follows. Solving of these sub-problems is much easier than the problem.

- Sub-problem I: Solving P^1
- Sub-problem II: Solving P with fixing X_{ij}^1 and Z_i^1 variables
- Sub-problem III: Solving P^1 with fixing Z_i^1 variables.

Other possible approaches to generate the neighborhood of this problem are to fix the network investment decisions or fix both facility location and network investment decisions. However, unlike the investigated approach, finding a feasible neighborhood solution with fixing the network decisions is too hard and usually leads to an infeasible solution that requires a great deal effort to repair it. On the other hand, reaching the optimal timing of network investment is quick and easy by fixing the facility location decisions. Therefore, at the first time period, the proposed algorithm uses simulated annealing to fix the facility location decisions and CPLEX is used to find the optimal decisions related to network investment. Then, CPLEX is used to solve the rest of the problem including the decisions regarding optimal facilities and network investments.

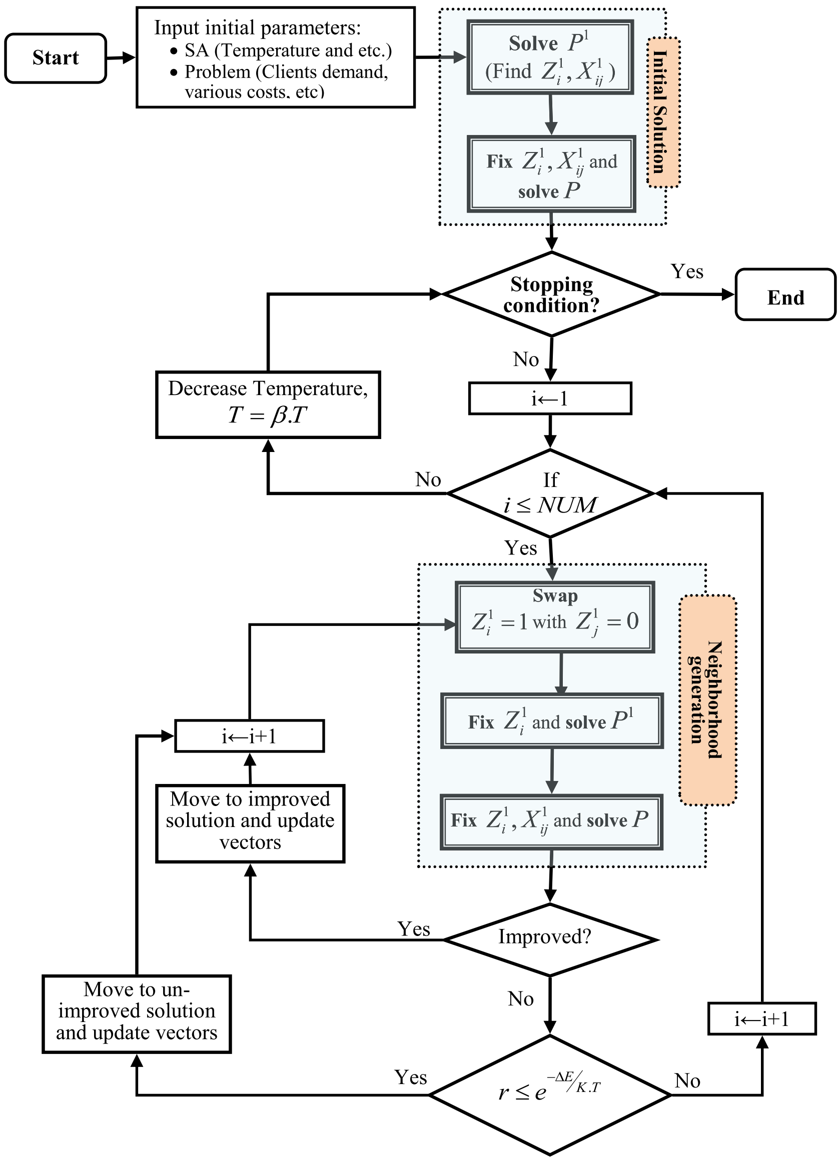

Fig. 3. The overall flowchart of the fix-and-optimize heuristic.

4. Experimental results

In this section, the results of a performed computational study are presented, and the performance of the proposed algorithms is also analyzed. The mixed-integer programming model described in the previous sections was solved with the CPLEX 12.1 solver and proposed heuristics on a range of randomly generated test problems and a specific real-world setting. The proposed algorithms were coded in Python 2.6. In addition, all programs were implemented on a dual quad core 2.66 GHz Intel Xeon X5550 processor with 32 GB of RAM.

4.1. Real-world application

A considerable amount of research in the facility location area has been advocated on implementing the models in the real-world applications. Regional planning is one of the special cases in developing countries where location decisions, commonly being made by local officials and based on political, economic, and cultural factors, can result in a placement of facilities that is far from optimal in terms of geographical accessibility [39]. Various studies [40–44] have shown the effectiveness of applying location analysis for health service development planning in developing nations. Two direct studies which explicitly addressed the problem of locating health facilities and network design were presented by Melkote [10] and

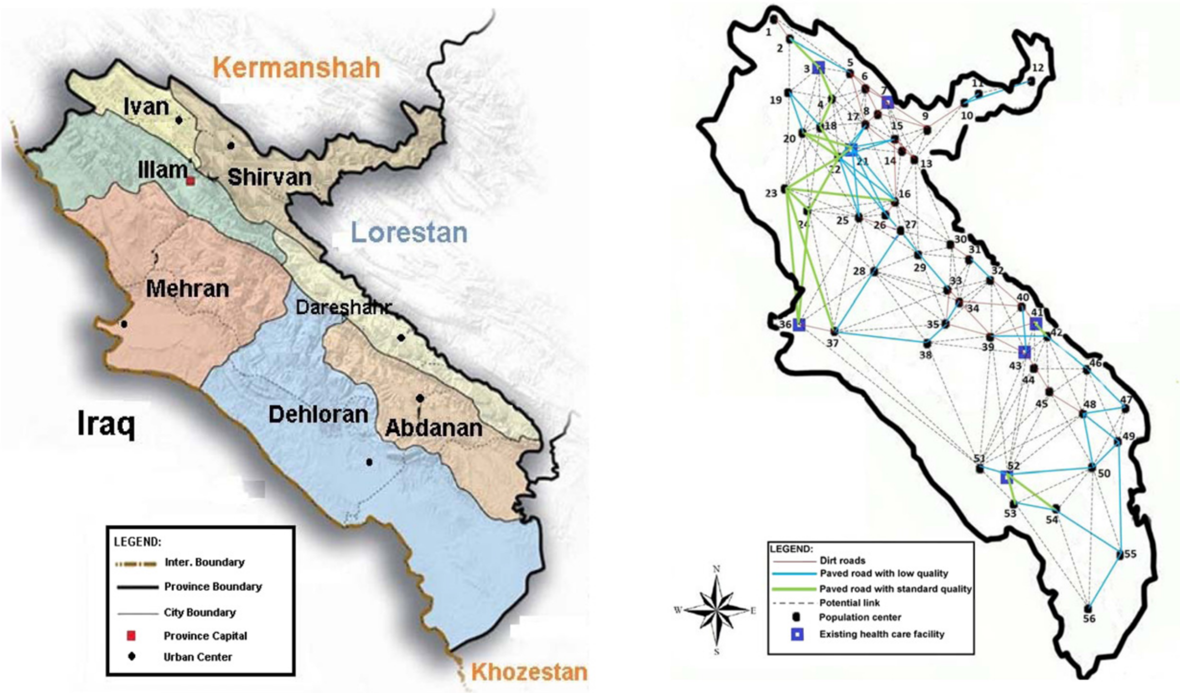

Fig. 4. Geographical map of Illam province with the urban residence centers and the road network (existing and potential).

Cocking [14]. Melkote [10] showed how the maximum covering FLNDP could be used in a real-life setting by applying it to the data taken from a regional planning problem in India. They compared the obtained results with a two-phase myopic approach of designing the network and locating facilities independently such that the potential saving using this model was quite significant. Furthermore, Cocking [14] demonstrated how the use of FLNDP analysis could be successfully integrated into health center planning of the Nouna health district in Burkina Faso to improve access to health facilities for the people.

A key point that should be considered in our model to implement in these kinds of applications is other benefits except health care accessibility. There are undoubtedly some other benefits that accrue to a community or region based on having more roads in addition to the health care benefits (e.g., accessibility to markets, to jobs, to educational institutes, to social and recreational areas, and to parts of the country beyond the immediate region). Furthermore, some qualitative factors such as tribal rivalries that sometimes influence which villages a person is willing to visit could be considered in the model. Hence, by looking only at the health care accessibility benefits, the model may well suggest investments in the network that are suboptimal when the full array of benefits associated with network upgrades are considered. Either of these considerations can be handled easily on the model using some modifications. By assuming e kind of service centers in the region, all parameters and variables except the link construction terms that means c_{ij}^t, h_{ij}^t and X_{ij}^t of the model are modified by adding an additional index. For example, in this case, the facility construction variable would be changed from Z_i^t to Z_i^{et} and defined as if a facility of kind e located on node i at time period t. Thus, the facilities' centers are located at the optimal sites and network decisions are made based on them.

Accessibility improvement to the health facilities for the rural population centers in the Illam Province of Iran is considered as an application of the problem under study. With the knowledge that there are some other benefits except health care in this region, the accessibility improvement to health care facilities in the studied region is so important and other benefits are ignored. This province is one of the deprived areas in the country that is divided into population districts for administration purposes. The Illam Province has 56 population centers that includes 19 cities and 37 rural districts with the city of Illam as the capital and its total population is about 550,000 [45]. Note that each rural district is composed of a number of villages, including a center. A geographical map, the population centers and existing network of our case study are depicted in the right side of Fig. 4.

There are some health facilities scattered throughout the district, including hospitals or large health centers at the urban centers, and other small health centers in rural districts, providing basic care. Fig. 4 shows the population centers numbered from 1 to 56 and the road network of Illam province, with those centers having a health facility indicated by rectangles. In the current conditions, there are 7 existing facilities and some network links that have been constructed, previously. These roads are classified into three categories in terms of quality. Dirt roads and paved roads in low and standard quality compose these categories and are numbered with 3, 2, and 1 respectively. Note that depending on the type of roads, transportation and operating costs vary during the planning horizon. In addition, the dirt and paved roads with low quality can be upgraded to standard roads with specific costs. These items might be considered as investment costs.

According to given explanations for our real-world application, some modification should be done on the proposed model. The roads with quality 2 and 3 can be upgraded to 1 with standard quality. Let $\hat{L}^t$ be the set of constructed links with rank 2 and 3. Moreover, $\hat{h}_{ij}^t$ and $\hat{r}_{ij}^{kt}$ represent the operating cost of link (i, j) and the transportation cost of client k on link (i, j) with rank 2 and 3 at time period t $((i, j) \in \hat{L}^t)$, respectively. Thus, based on these definitions, the objective function of the model can be modified as follows:

$$\textbf{Min} \quad \sum_{t\in T}\sum_{(i,j)\in L^t\setminus\hat{L}^t}\sum_{k\in N} r_{ij}^{kt}Y_{ij}^{kt} + \sum_{t\in T}\sum_{(i,j)\in\hat{L}^t}\sum_{k\in N}[\hat{r}_{ij}^{kt}(1-X_{ij}^t) + r_{ij}^{kt}X_{ij}^t]Y_{ij}^{kt}$$
$$+ \sum_{t\in T}\sum_{i\in N} f_i^t Z_i^t + \sum_{t\in T}\sum_{(i,j)\in L^t\setminus\hat{L}^t:i<j} h_{ij}^t X_{ij}^t + \sum_{t\in T}\sum_{(i,j)\in\hat{L}^t:i<j}[\hat{h}_{ij}^t(1-X_{ij}^t) + h_{ij}^t X_{ij}^t]. \tag{25}$$

In addition, $X_{ij}^0 = 1$, should be only held for the roads with rank 1 or standard quality and constraint (5) is no longer needed for $(i, j) \in \hat{L}^t$. For the sake of simplicity, we assume $\hat{r}_{ij}^{kt} = r_{ij}^{kt}$ in our considered application.

In Iran, the ministry of health and medical education and the ministry of road and transportation are responsible for investment in health facilities and road network construction, respectively. Hence, these two different organizations provide the budget for each year during a planning horizon. In addition, the unspent budget at the end of time periods is not returned and remains for the next time period. As a result, according to this description, our formulation is appropriate to the model of the application at hand.

To improve the physical access from the people of Illam province to the health facilities within the district, the situation was investigated as a budget-constrained dynamic facility location–network design problem. Hence, a new facility at any node that does not already have one may be built. In addition, any of a number of different roads may be constructed or upgraded. Problem options for improvement include building new facilities, constructing new roads and upgrading the roads with low quality (level 2 and 3) to the standard quality (level 1).

As far as possible, reliable data was collected for the problem. There are 7 existing facilities and 49 potential nodes to open a facility on them as well as 87 existing and 113 potential links. These roads have three different qualities and shown with different thicknesses in the graph so that they can be distinguished from each other. Every population center is a client node with demand equal to its population. For each subsequent period, the population is increased due to the growth rate in the rural and urban centers as 1.2% and 2.9%, respectively [45]. The fixed cost of opening a facility depends on the node's demand and varies between 1018 and 3749 for the first period. In addition, the operating cost of the facility for the first time period was considered as at least 10% and at most 12% of the opening cost of the facility. The construction cost of a new road is calculated as 100 per kilometer, and annual operating cost is considered as 2%, 3% and 4% of this cost for different types of roads according to their qualities. The traveling cost for each client in kilometer is randomly calculated subject to a discrete uniform distribution in [0.0003, 0.0004]. Moreover, for each subsequent period the value of mentioned parameters was increased by an inflation rate equal to 10%.

Fig. 5, presents the optimal solution of the problem for the existing network at the first and second time period. At the first period, subject to the available budget for facilities and links that respectively equal 1800 and 7200, one health facility is opened on node 26 and one new link is also constructed between nodes 25 and 26. In addition, two existing dirt roads between nodes (14, 17) and (30, 31) are upgraded to the road with the standard quality (level 1). The moving direction of customers to health care facilities is shown on links and the set of population centers that are serviced by a facility is also shown with a hull in all networks. For example, at the first time period, the people in population center 12 should be serviced by a located health care on node 7 through route (12-10-9-7). In addition, people in population centers (5, 6, 7, 8, 9, 10, 11, 12) are serviced by a located health care on node 7. If one compares the existing road system (Fig. 4) to the solution in Fig. 5, it is easy to see that there are some roads that are not used by people for health care accessibility at the optimal solution. As mentioned before, these roads may be used for the other purposes like accessibility to markets and to jobs, etc. However, we minimize the total costs, including: transportation cost and operating cost of facilities and links in the proposed model. Thus, some roads like (14, 17) at the first time period despite not being used for health care accessibility purposes, may be upgraded.

By opening new facilities on some nodes and investment on links over time, some flow and allocation variables may also be changed. For example, at the first time period, people living in population center 29 are serviced by the located health care on node 26 and travel through route (29-27-26). At the second time period, a new facility is opened on node 33 that services people in 29. Furthermore, a number of other people near this facility that traveled to other facilities at the first time period, are also serviced by this facility. Unlike other time periods, there is no investment at the fourth time period on facilities and links and the optimal network for this period is the same as the third period shown in Fig. 6.

Table 2 shows the cost and budget data of the application for 5 considered time periods and the entire solution. From this table, we can readily see that the operating cost of facilities and roads is increased during the planning horizon. In addition, decreasing transportation costs at the third time period is due to opening a facility on nodes (11, 54), simultaneously. The total available budget for investment in facilities and roads is 6600 and 26,400, respectively that is divided between 5 time periods. During these time periods, four new facilities and two links are constructed and 6 dirt existing roads are also upgraded to the standard level with the considered budget. Certainly, in addition to issues of cost, increasing the accessibility to health care facilities and eradication of poverty can be one of the basic results of implementing such systems.

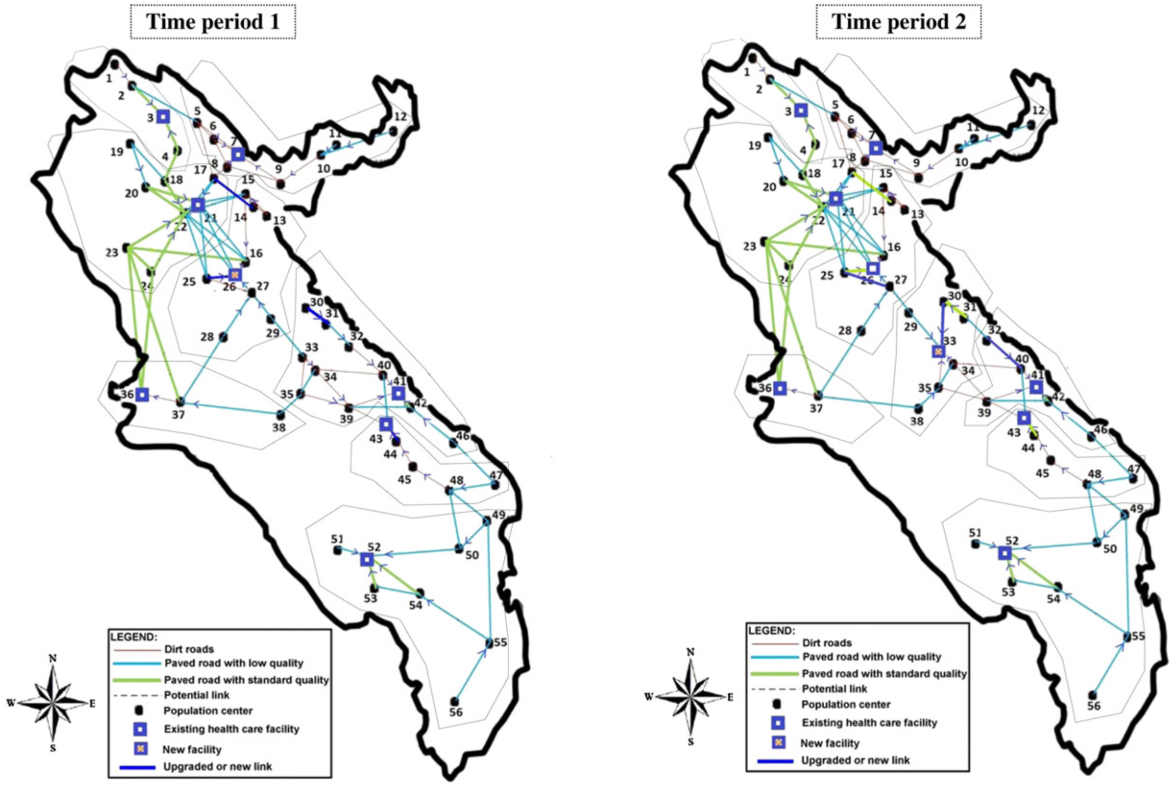

Fig. 5. The optimal solution of the problem for the existing application at the first and second time period.

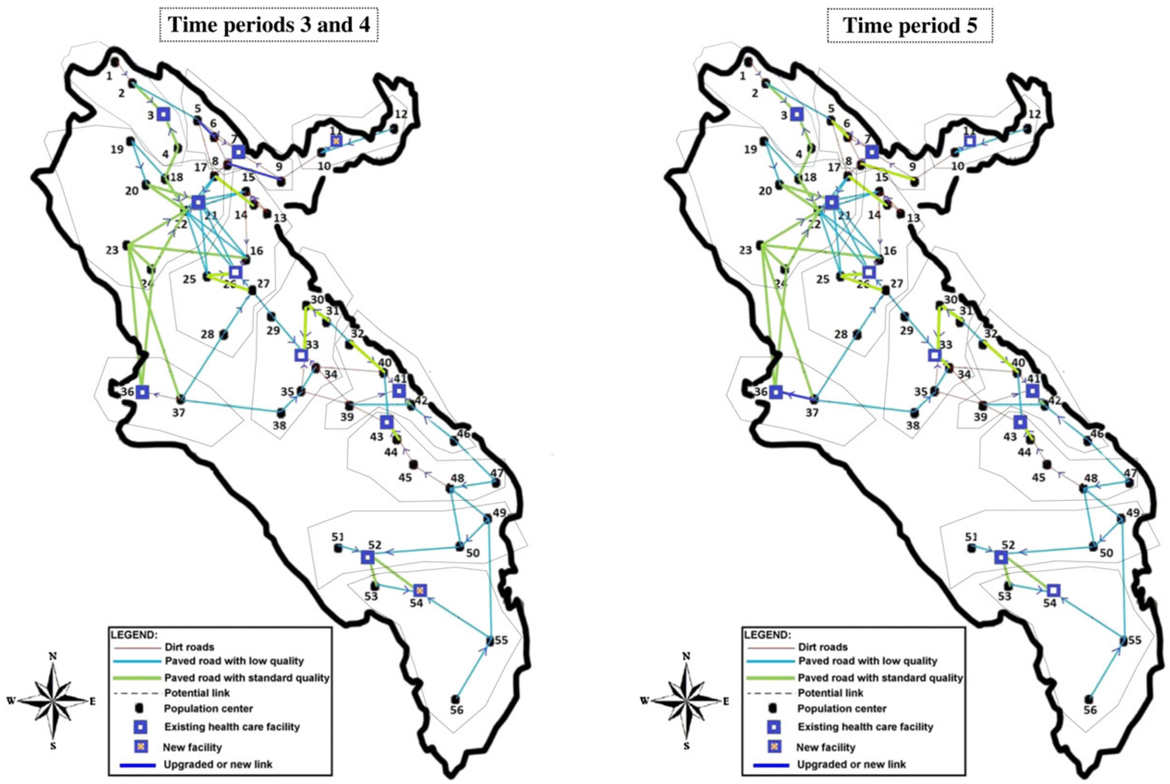

Fig. 6. The optimal solution of the problem for the existing application at the first and second time period.

Table 2
The cost and budget data of the application for 5 considered time periods and the entire problem.

Time period	Costs				Budget			
	Traveling	Operating		Total	Available		Invested	
		Facility	Road		Facilities	Roads	Facilities	Roads
$T = 1$	794.75	2 620.12	12 653.41	16 068.28	1800	7 200	1627.97	7 162.00
$T = 2$	873.33	2 983.73	13 972.40	17 829.46	1800	7 200	1231.66	7 231.68
$T = 3$	637.93	3 646.31	15 623.47	19 907.71	2000	8 000	2718.51	7 999.17
$T = 4$	701.89	3 900.83	16 873.35	21 476.07	0	0	0.00	0.00
$T = 5$	772.35	4 305.49	18 489.87	23 567.71	1000	4 000	0.00	3 999.87
Total	**3780.25**	**17 456.48**	**77 612.5**	**98 849.23**	**6600**	**26 400**	**5578.14**	**26 392.72**

Table 3
Parameters used in the random generation of instances.

Symbol	Description	Value tested
d_k^t	The demand of client k during time period t	$d_k^t = d_k^{t-1} \cdot (1 + U[0, 5]\%),\ t \geq 2$ $d_k^1 = U[10, 100]$
g_i^t	Opening cost of facility	$g_i^t = g_i^{t-1} \cdot (1 + U[2, 10]\%),\ t \geq 2$ $g_i^1 = U[1200, 1500]$
f_i^t	Operating cost of facility	$\delta.g_i^t,\ \delta = 0.06$
ρ_{ij}^t	Travel cost of flow unit on link (i, j) at time period t	$\rho_{ij}^t = (1+U[0, 5]\%) \cdot \rho_{ij}^{(t-1)},\ t \geq 2,$ $\rho_{ij}^1 = \varepsilon \cdot m_{ij},\ \varepsilon = 0.01$
r_{ij}^{kt}	Travel cost of client k on link (i, j) at time period t	$r_{ij}^{kt} = \rho_{ij}^t d_k^t$
h_{ij}^t	Construction cost of link (i, j) at time period t	$h_{ij}^t = (1 + U[0, 5]\%) \cdot h_{ij}^{t-1},\ t \geq 2,$ $h_{ij}^1 = \sigma \cdot m_{ij},\ \sigma = 0.1$
c_{ij}^t	Operating cost of link (i, j) at time period t	$c_{ij}^t = \upsilon \cdot h_{ij}^t,\ \upsilon = 0.05$

4.2. Generated instances

Besides the above application, many test problems have been generated for the model. The dimensions of test problems vary from small to large scale. The test problems were solved with a number of network nodes from 20 to 80 and time periods from 5 to 20. The results obtained show that the performance of CPLEX is different for these test problems. CPLEX was unable to optimally solve the instances in a reasonable execution time. It could solve some small test problems but with much CPU time and relatively large GAP. Furthermore, CPLEX did not solve most of the instances in a given CPU time. As a result, the time needed to obtain a feasible or good solution of the proposed model by using CPLEX is very dependent on the dimensions of the problem. Thus, while solving the model, the CPU time was limited to $(50|N|\,|T|)$ s on CPLEX. For example, the CPU time for a test problem with 60 nodes and 20 time periods was 60 000 s, which is more than 16 h.

4.2.1. Data generation

Some test problems were generated randomly with the following process. The following features are valid for all test problems:

– The locations of the clients (network nodes) are generated randomly and uniformly distributed over a $100 * 100$ area.
– The construction cost of a link is proportional to the length of the link.
– The link travel cost is proportional to the length of the link.
– The operating cost of a link is considered as a percentage of the construction cost of the link.

Table 3 gives the details relating to the generation of demand, costs and other input data. In this table, $U[a, b]$ denotes the random generation of numbers in the interval $[a, b]$ according to a discrete uniform distribution. In the period one, entries are drawn from a discrete uniform distribution. Then, in each subsequent period $t \geq 2$, the generated value changes by a percentage that varies between $c\%$ and $d\%$. For instance, having d_k^t with considered value means that the demand of client k during the first period is randomly generated subject to a discrete uniform distribution in [10, 100]. The demand in the second period is changed or increased at most 5% compared to the first period.

The following process is considered to predict the budget parameter in new networks at each time period. Melkote [10] in his thesis showed that for the static case if α facilities are opened in a feasible solution, then $(N - \alpha)$ links should be constructed. Thus, this property is used here to generate the budget parameter at the first time period. We assume that $\beta \cdot N$ facilities and $(1 - \beta) \cdot N$ links are opened at the first time period; hence, $(1 - \beta) \cdot N$ links are constructed. By multiplying

Table 4
The CPLEX performance to solve the sub-problems I and II for the new network.

Test problems	The number of parameters			CPLEX solver	
	N	L	T	CPU time (s) (subproblem I)	CPU time (s) (subproblem II)
TP1N	20	46	5	421	6
TP2N	20	46	10	344	27
TP3N	20	46	20	445	157
TP4N	20	61	5	1284	6
TP5N	20	61	10	1187	21
TP6N	20	61	20	1181	62
TP7N	40	137	5	456	63
TP8N	40	137	10	450	2846
TP9N	40	137	20	454	2754
TP10N	40	162	5	296	99
TP11N	40	162	10	303	1519
TP12N	40	162	20	289	10227

on the average fixed cost of facilities and links, the following value is obtained. In our generated data, the value of β was set to 0.05.

$$\gamma_1 = \beta \cdot N \cdot \bar{g}_1 \tag{26}$$

$$\eta_1 = (1-\beta) \cdot N \cdot \bar{h}_1 \tag{27}$$

that:

- $\bar{g}_t$: is the average value of facilities fixed cost at time period t
- h_t: is the average value of links fixed cost at time period t.

Now, the maximum budget for facilities and links at the first time period are selected from an interval $[0.8\gamma_1, 1.2\gamma_1]$ and $[0.8\eta_1, 1.2\eta_1]$, respectively. For the other time periods, the average ratio of increasing facility and link fixed costs is calculated and multiplied in the budget at the first time period.

$$\gamma_t = (\bar{g}_t/\bar{g}_1) \cdot \bar{B}_1 \tag{28}$$

$$\eta_t = (\bar{h}_t/\bar{h}_1) \cdot \hat{B}_1. \tag{29}$$

Then, the budget for facilities and links at time period ($t \geq 2$) is generated in interval $[0, 0.4\gamma_t]$ and $[0, 0.2\eta_t]$.

4.3. Computational results

Twenty seven test problems with various dimensions were solved by the proposed heuristics and the CPLEX solver to construct a new network and improve an existing network. The instances characteristics with the obtained results are shown in Tables 4–6. The first column in these tables indicates the instance such that the letters "*N*" and "*E*" specify the kind of network as new and existing networks, respectively. The budget parameter for the existing networks is assumed half of calculated value for the new networks by Eqs. (26)–(29). All other data are assumed similar. CPLEX solver could solve optimally most of the instances for networks with some existing facilities and links. Thus, the proposed algorithms were not used to the problems with existing networks. In addition, in new networks, the proposed heuristics results are compared with the lower and upper bounds obtained by CPLEX solver. The GAP of the CPLEX solver and the proposed heuristics is computed as follows. *Obj.* is the value of the optimum or the best found solution, and *LB* is the lower bound reported by CPLEX after the given time limit.

$$\text{GAP} = \frac{\text{Obj.} - \text{LB}}{\text{Obj.}} * 100. \tag{30}$$

The criterion for the termination of the proposed algorithms and CPLEX is one of the following conditions:

(1) A specified time limit has been reached (set as $SP = 50|N|\,|T|$).
(2) The gap between the lower bound and upper bound is zero.

The computational complexity of the model depends significantly on the budget constraints. Solving the model would be much easier, if we relax all or part of these constraints. These circumstances have seen in the considered sub-problems in the proposed heuristics. Table 4 gives the results obtained for an experiment aimed at studying the performance of the CPLEX to solve the subproblems I and II at the new networks. The CPLEX solver could solve all instances in this table optimally in a small running time requirement. We do not report the results for sub-problem III, since it is a special case of subproblems I when Z_i^1 variables are fixed.

Table 5
The number of instances parameters for the existing networks and computational results of CPLEX.

Test problems		The number of parameters					CPLEX			
		N	*L*	*T*	Existing facilities	Existing links	Obj.	*LB*	Gap (%)	CPU time (s)
TP1E		20	46	5	3	15	14 924.87	14 924.87	0.00	4
TP2E		20	46	10	3	15	29 698.23	29 698.23	0.00	14
TP3E		20	46	20	3	15	67 676.54	67 676.54	0.00	81
TP4E		20	61	5	3	14	31 419.98	31 419.98	0.00	5
TP5E		20	61	10	3	14	67 884.16	67 884.16	0.00	17
TP6E		20	61	20	3	14	158 091.31	158 091.31	0.00	77
TP7E		40	137	5	4	36	21 159.56	21 159.56	0.00	63
TP8E		40	137	10	4	36	45 581.23	45 581.23	0.00	1 758
TP9E		40	137	20	4	36	111 643.23	110 830.16	0.73	40 000
TP10E		40	162	5	2	38	19 232.89	19 232.89	0.00	289
TP11E		40	162	10	2	38	40 310.81	39 730.11	1.44	20 000
TP12E		40	162	20	2	38	99 103.48	96 797.34	2.33	40 000
TP13E		60	180	5	8	52	25 110.09	25 110.09	0.00	347
TP14E		60	180	10	8	52	56 892.74	56 892.74	0.00	1 479
TP15E		60	180	20	8	52	151 115.62	149 518.50	1.06	60 000
TP16E		60	205	5	5	55	26 973.29	26 973.29	0.00	480
TP17E		60	205	10	5	55	58 636.16	58 636.16	0.00	16 514
TP18E		60	205	20	5	55	144 200.68	142 289.95	1.33	60 000
TP19E		80	171	5	7	65	33 696.53	33 696.53	0.00	5 622
TP20E		80	171	10	7	65	74 677.35	73 256.40	1.90	40 000
TP21E		80	171	20	7	65	198 074.52	192 701.54	2.71	80 000
TP22E		80	280	5	8	72	30 999.44	30 999.44	0.00	1 179
TP23E		80	280	10	8	72	70 092.80	69 623.72	0.67	40 000
TP24E		80	280	20	8	72	183 511.97	179 939.71	1.95	80 000
	TP25E	56	200	5	7	87	98 750.91	98 750.91	0.00	625
Case study	TP26E	56	200	10	7	87	236 530.52	236 461.00	0.03	28 000
	TP27E	56	200	20	7	87	741 575.22	740 644.69	0.13	56 000
Average							**105 094.97**	**104 389.67**	**0.59**	

4.3.1. Improving an existing network

In this sub-section, our aim is to improve service to clients through new facilities and links while considering the existing facilities and links. As mentioned before, solving the model with some constructed facilities and links is not very difficult with CPLEX. The results for this set of experiments are summarized in Table 5. In this table, column 1 indicates each instance. The number of parameters for each instance is shown in columns 2–4. The number of existing facilities and network links are represented in columns 5 and 6. The obtained results from CPLEX are given in the next four columns; where Obj. is the best found integer solution at the time of termination (CPU time), LB is the lower bound and Gap is the relative gap between lower and upper bound. A larger GAP usually indicates that CPLEX would require longer run times to converge or terminate. CPLEX achieved the optimal solution for 60% of instances and found a pretty good solution for the other instances with a small integrality gap. As expected, the performance of CPLEX is directly dependent on the problem dimensions.

4.3.2. Constructing a new network

Table 6 presents the results obtained from CPLEX, Heuristic, and hybrid SA heuristic for twenty seven instances in new networks. The number of instance parameters with the CPLEX results is given in columns 2–8. The results of proposed heuristics are reported in the other columns. The Time(s) column reports the running time that the heuristic reached to the best solution or Obj. In addition, the gap ratio of the proposed heuristics vs. CPLEX is demonstrated in columns with label Ratio.

Note that in this category, the performance of CPLEX is so bad that it even could not handle small-size instances such that for the majority of instances, CPLEX could not even find a feasible solution. Hence, unlike improving the existing networks, the CPLEX solver does not provide a suitable tool to solve the problem with the networks without any existing facility or links and using heuristics becomes necessary. In order to implement the proposed hybrid SA algorithm efficiently, the following parameters are used. The initial temperature was set as $0.01OF_1$ such that $OF1$ is the value of the objective function for the initial solution. The temperature updating scheme is by $T_{i+1} = 0.9T_i$. The number of iterations in each temperature is NUM $= 10$. One of the main parameters used during the proposed heuristics is the running time of CPLEX to solve the sub-problems. As mentioned before three sub-problems are used in the proposed algorithms. For the first heuristic, the sub-problem (I) will break after $0.5SP$ if it fails to reach the optimal solution. The remaining time is used to solve the sub-problem (II). In addition, at the hybrid SA, the CPLEX running time constraint for the sub-problems (I), (II) and (III) is 0.2, 0.05, and 0.03 of stopping time criterion (SP), respectively.

Table 6
The number of test problems parameters for new networks and computational results of CPLEX and proposed heuristics.

Test problems		The number of parameters			CPLEX				Heuristic				Hybrid SA			
		N	*L*	*T*	Obj.	*LB*	Gap (%)	CPU time (s)	Obj.	Gap (%)	Time (s)	Ratio	Obj.	Gap (%)	Time (s)	Ratio
TP1N		20	46	5	20 558.72	19 115.27	7.02	5 000	20 473.40	**6.63**	427	0.94	20 473.40	**6.63**	427	0.94
TP2N		20	46	10	NA	32 703.29	–	10 000	35 739.84	8.50	371	–	35 528.03	**7.95**	397	–
TP3N		20	46	20	NA	67 838.39	–	20 000	71 761.71	5.47	604	–	70 973.67	**4.42**	645	–
TP4N		20	61	5	36 115.31	32 264.09	10.66	5 000	36 425.75	11.43	1 289	1.07	35 614.49	**9.41**	1 455	0.88
TP5N		20	61	10	70 431.56	57 654.74	18.14	10 000	70 202.49	17.87	1 208	0.99	69 581.22	**17.14**	1 534	0.94
TP6N		20	61	20	NA	127 629.86	–	20 000	152 077.40	16.08	1 869	–	151 545.64	**15.78**	4 497	–
TP7N		40	137	5	23 807.35	21 651.39	9.06	10 000	23 120.55	**6.35**	519	0.70	23 120.55	**6.35**	519	0.70
TP8N		40	137	10	NA	43 284.98	–	20 000	46 256.15	**6.42**	2 296	–	46 256.15	**6.42**	2 296	–
TP9N		40	137	20	NA	100 874.01	–	40 000	105 190.16	4.10	3 238	–	105 164.48	**4.08**	6 501	–
TP10N		40	162	5	NA	18 430.57	–	10 000	20 451.44	9.88	396	–	19 964.54	**7.68**	2 834	–
TP11N		40	162	10	39 092.04	35 655.01	8.79	20 000	38 884.58	8.31	1 822	0.94	37 415.17	**4.70**	2 605	0.54
TP12N		40	162	20	89 971.40	87 563.42	**2.68**	40 000	92 249.74	5.08	2 521	1.90	90 390.87	3.13	9 828	1.17
TP13N		60	180	5	30 804.62	27 901.62	9.42	15 000	28 867.19	**3.34**	7 428	0.35	28 941.91	3.59	1 758	0.38
TP14N		60	180	10	NA	57 757.15	–	30 000	60 742.95	4.92	14 627	–	60 090.00	**3.88**	26 262	–
TP15N		60	180	20	NA	142 325.99	–	60 000	149 982.60	5.11	60 000	–	146 898.69	**3.11**	15 819	–
TP16N		60	205	5	NA	26 336.77	–	15 000	28 466.98	7.48	7 782	–	27 894.80	**5.59**	8 252	–
TP17N		60	205	10	NA	53 736.60	–	30 000	56 908.84	5.57	17 136	–	56 412.09	**4.74**	21 361	–
TP18N		60	205	20	139 262.41	130 562.90	6.25	60 000	135 622.64	3.73	60 000	0.60	134 023.57	**2.58**	13 660	0.41
TP19N		80	171	5	44 652.63	36 068.30	19.22	20 000	37 832.36	4.66	10 805	0.24	37 577.65	**4.02**	19 255	0.21
TP20N		80	171	10	NA	73 508.12	–	40 000	76 229.48	**3.57**	40 000	–	76 343.88	3.71	15 477	–
TP21N		80	171	20	NA	178 843.89	–	80 000	187 581.87	4.66	80 000	–	185 716.07	**3.70**	24 659	–
TP22N		80	280	5	34 598.10	32 321.61	6.58	20 000	33 957.52	4.82	8 332	0.73	33 504.38	**3.53**	12 874	0.54
TP23N		80	280	10	74 391.75	68 562.54	7.84	40 000	71 706.03	4.38	40 000	0.56	71 098.73	**3.57**	31 868	0.46
TP24N		80	280	20	NA	168 507.11	–	80 000	175 401.41	3.93	80 000	–	173 903.28	**3.10**	60 596	–
	TP25N	56	200	5	NA	45 805.74	–	14 000	51 321.10	**10.75**	7 169	–	51 321.10	**10.75**	5 815	–
Case study	TP26N	56	200	10	NA	111 870.74	–	28 000	121 996.75	8.30	14 696	–	121 205.52	**7.70**	12 424	–
	TP27N	56	200	20	NA	374 135.96	–	56 000	400 335.43	**6.54**	32 439	–	400 335.43	**6.54**	28 504	–
Average					–	**80 478.15**	–	29 556		**6.96**	18 406			**6.07**	12 300	

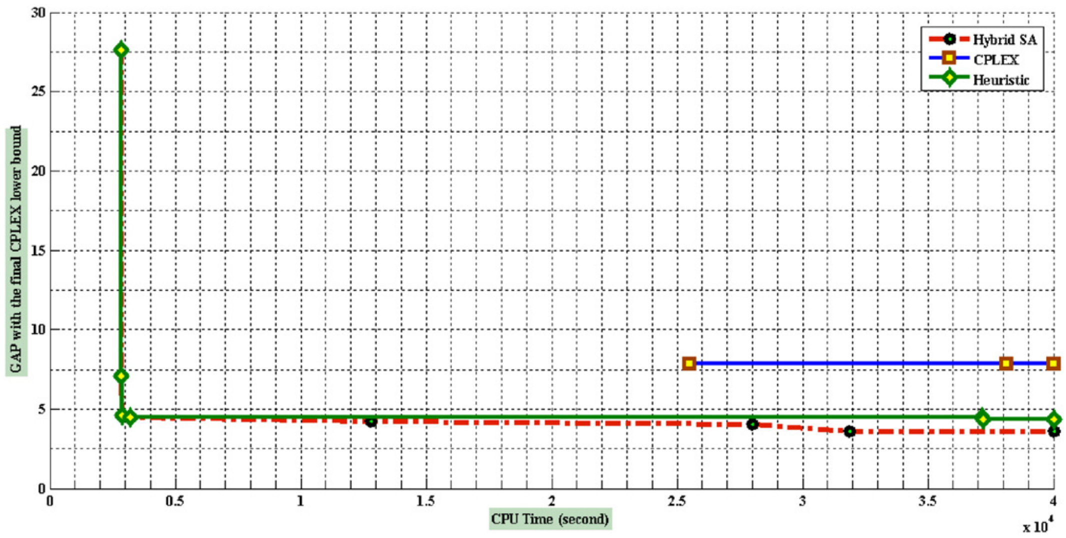

Fig. 7. Behavior of the CPLEX, heuristic and the hybrid SA heuristic to solve TP23 over time.

The CPLEX solver requires an extraordinarily long running time to find a feasible solution for most instances. In addition, as shown in Table 6, the CPLEX failed to find a feasible solution for nearly 60% of instances in the considered computation time. On the other hand, the heuristics reach a feasible solution for all instances in a small computation time. Moreover, for the instances where CPLEX found a feasible solution, none of the solutions were reported to be optimal. For the majority of instances, the quality of the solution found by the proposed heuristics taking much less time outperformed that of the CPLEX solver taking a time in the order of thousands of seconds. These results reveal that the GAP ratio of the proposed heuristics vs. CPLEX is less than one for all test problems except for the three ones. Thus, CPLEX only has better performance than the proposed heuristics for these instances. This experimental study clearly indicates the advantages of using the proposed heuristics in solving the budget-constrained DUFLNDP.

For most instances, the results from the hybrid SA are much better than the results obtained from the greedy heuristic. Taking all instances together, the difference in average deviation of the solution obtained by the hybrid SA algorithm from the heuristic with the lower bound of CPLEX is almost 0.9% (6.96–6.07) (last row).

Note that the relatively high value of GAP is due to the large difference in the lower bound reported by CPLEX and the optimal solution. To clarify this issue and the computational results, test problems 4 and 5 were selected and solved by CPLEX with running time equal to $200|N|\,|T|$ s. After this time, CPLEX could not reach the optimal solution, and the obtained value of the objective function for these two test problems was 36 115.31 and 70 416.17 with the lower bound 34 342.20 and 61 287.79, respectively. Therefore, by increasing the running time for these instances, the lower bound of the problem was improved and their best integer solution did not have a significant improvement. CPLEX has a very smooth trend after a specified time. Thus, proposed heuristics show good performance in comparison to CPLEX. Consequently, these results show that the heuristic and hybrid SA generate much better solutions with less time than the CPLEX solver, and therefore are capable of handling larger problems. Hence, the development of these heuristics appears to be worthwhile.

The behavior of CPLEX, heuristic and the fix-and-optimize heuristic were also studied for solving the model over time. For this study, test problem 23 was selected, and its results are plotted in Fig. 7. As expected, by increasing the execution time, the calculated GAP of the best found solution decreases in all approaches. It is interesting to note that CPLEX finds the first feasible solution at a time of 25 502 s with a gap of 7.84%. However, the heuristic and hybrid SA start with a solution with a gap of 25.24% at time 2850 s and immediately reach a very good solution with Gap 4.62 after 12 s from the initial solution. Furthermore, with the run time limit, CPLEX can slightly improve the solution only one time with 0.01% change. On the other hand, heuristic and the hybrid SA improve the current solution 5 and 7 times, respectively. As shown in Fig. 7, CPLEX spends too much time to reach a feasible solution.

5. Conclusions

This paper presented an optimization mixed-integer linear programming model for the budget-constrained dynamic (multi-period) uncapacitated facility location–network design problem. The problem is aimed at the determination of facility locations required to satisfy a set of clients' demands and the determination of travelable links to connect clients to facilities. Furthermore, assigning clients to open facilities during the planning horizon is also determined by the model. Based on the number of organizations involved in investment and their policies during the planning horizon, an approach was employed to consider the budget constraint for opening the facilities and constructing the network links. As an application of the problem at hand and for the model, the accessibility of health facilities in Illam Province of Iran was investigated. The usefulness of budget-constrained DUFLNDP results as a decision-making aid in this real-world context was demonstrated. Furthermore, a greedy heuristic and a fix-and-optimize heuristic based on SA and exact methods (Branch & Bound and cutting methods) were proposed to solve the model. In these algorithms, the problem is split into some easier subproblems. In addition, at the hybrid SA, the binary variables of facilities opening at the first period are set to a fixed value which these binary variables are iteratively determined based on simulated annealing algorithm. Then, a series of subproblems that are derived from the budget-constrained DUFLNDP are solved iteratively and quickly with CPLEX solver. The performance of the proposed heuristics was compared with that of the CPLEX solver in the numerical experiments. The results showed that the heuristic and the hybrid SA produce much better results than CPLEX.

Areas of further research are developing algorithms to solve large-scale instances for the proposed models. Further attention is also required in the future to include additional real assumptions such as uncertainty and capacity of the facilities or/and the network for the problem. Another line of research would be enhancing the model to take into account accessibility to different classes of facilities with various levels of hierarchy (see Bigotte et al. [16]).

Acknowledgments

We would like to thank the anonymous referee for his/her constructive comments which have strongly improved the insights and presentation of the paper. We also acknowledge Mr. David Todorovski for the invaluable assistance in the English editing of the paper.

References

[1] A. Weber, Uber den Standort der Industrian, University of Chicago Press, 1909, Translated as (in 1929): Alfred Weber's Theory of the Location of Industries.
[2] A. Klose, A. Drexl, Facility location models for distribution system design, European Journal of Operational Research 162 (2005) 4–29.
[3] G. Nagy, S. Salhi, Location-routing: issues, models and methods, European Journal of Operational Research 177 (2007) 649–672.
[4] S. Alumur, B.Y. Kara, Network hub location problems: the state of the art, European Journal of Operational Research 190 (2008) 1–21.
[5] C.S. ReVelle, H.A. Eiselt, M.S. Daskin, A bibliography for some fundamental problem categories in discrete location science, European Journal of Operational Research 184 (2008) 817–848.
[6] M.T. Melo, S. Nickel, F. Saldanha-da-Gama, Facility location and supply chain management—a review, European Journal of Operational Research 196 (2009) 401–412.
[7] T.L. Magnanti, R.T. Wong, Network design and transportation planning: models and algorithms, Transportation Science 18 (1984) 1–55.
[8] H. Yang, M. Bell, Models and algorithms for road network design: a review and some new developments, Transport Reviews 18 (1998) 257–278.
[9] M.S. Daskin, A.P. Hurter, M.G. VanBuer, Toward an integrated model of facility location and transportation network design, Working Paper, Transportation Center, Northwestern University, 1993.
[10] S. Melkote, Integrated models of facility location and network design, Industrial Engineering and Management Science, Northwestern University, Doctor of Philosophy, Evaston, Illinois, 1996.
[11] S. Melkote, M.S. Daskin, An integrated model of facility location and transportation network design, Transportation Research, Part A 35 (2001) 515–538.
[12] S. Melkote, M.S. Daskin, Capacitated facility location–network design problems, European Journal of Operational Research 129 (2001) 481–495.
[13] Z. Drezner, G.O. Wesolowsky, Network design: selection and design of links and facility location, Transportation Research, Part A 37 (2003) 241–256.
[14] C. Cocking, Solutions to facility location–network design problems, Department of Computer Science, University of Heidelberg, Doctor of Philosophy Thesis, 2008.
[15] C. Cocking, G. Reinelt, Heuristics for budget facility location–network design problems with minisum objective, in: Operations Research Proceedings 2008, Springer, Berlin, Heidelberg, 2009, pp. 563–568.
[16] J.F. Bigotte, D. Krass, A.P. Antunes, O. Berman, Integrated modeling of urban hierarchy and transportation network planning, Transportation Research, Part A 44 (2010) 506–522.
[17] J. Current, S. Ratick, C. ReVelle, Dynamic facility location when the total number of facilities is uncertain: a decision analysis approach, European Journal of Operational Research 110 (1997) 597–609.
[18] A.S. Manne, Investment for Capacity Expansion: Size, Location, and Time Phasing, MIT Press, Cambridge, MA, 1967.
[19] A.S. Manne, Capacity expansion and probabilistic growth, Econometrica 29 (1961) 632–649.
[20] R.H. Ballou, Dynamic warehouse location analysis, Journal of Marketing Research 5 (1968) 271–276.
[21] R.Z. Farahani, Z. Drezner, N. Asgari, Single facility location and relocation problem with time dependent weights and discrete planning horizon, Annals of Operations Research 167 (2009) 353–368.
[22] J. Dias, M. Eugénia Captivo, J. Clímaco, Efficient primal–dual heuristic for a dynamic location problem, Computers & Operations Research 34 (2007) 1800–1823.
[23] A.J. Scott, Dynamic location-allocation systems: some basic planning strategies, Environment and Planning 3 (1971) 73–82.
[24] Z. Drezner, Dynamic facility location: the progressive p-median problem, Location Science 3 (1995) 1–7.
[25] J.A. Mesa, Multi-period medians on networks, Operations Research 25 (1991) 87–95.
[26] J.W. Chrissis, R.P. Davis, D.M. Miller, The dynamic set covering problem, Applied Mathematical Modelling 6 (1982) 2–6.
[27] G. Gunawardane, Dynamic versions of set covering type public facility location problems, European Journal of Operational Research 10 (1982) 190–195.
[28] G. Laporte, P. Dejax, Dynamic location-routing problems, Journal of the Operational Research Society 40 (1989) 471–482.
[29] M. Albareda-Sambola, A. Alonso-Ayuso, F.L. Escudero, E. Fernández, Y. Hinojosa, C. Pizarro-Romero, A computational comparison of several formulations for the multi-period incremental service facility location problem, TOP 18 (2009) 62–80.
[30] Y. Hinojosa, J. Puerto, F. Fernández, A multi-period two-echelon multicommodity capacitated plant location problem, European Journal of Operational Research 123 (2000) 271–291.
[31] J.E.T. Soto, Dynamic and robust capacitated facility location in time varying demand environments, in: Industrial and Systems Engineering, Texas A & M University, 2009.
[32] R.Z. Farahani, M. Abedian, S. Sharahi, Dynamic facility location problem, in: Facility Location: Concepts, Models, Algorithms and Case Studies, Springer, 2009.
[33] D.H. Lee, M. Dong, Dynamic network design for reverse logistics operations under uncertainty, Transportation Research Part E 45 (2009) 61–71.
[34] S.V. Ukkusuri, G. Patil, Multi-period transportation network design under demand uncertainty, Transportation Research Part B 43 (2009) 625–642.
[35] Y. Pochet, L.A. Wolsey, Production Planning by Mixed Integer Programming, Springer, New York, NY, 2006.

[36] S. Helber, F. Sahling, A fix-and-optimize approach for the multi-level capacitated lot sizing problem, International Journal of Production Economics 123 (2010) 247–256.
[37] F. Sahling, L. Busch, H. Tempelmeier, S. Helber, Solving a multi-level capacitated lot sizing problem with multi-period setup carry-over via a fix-and-optimize heuristic, Computers & Operations Research 36 (2009) 2546–2553.
[38] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science 220 (1983) 621–630.
[39] C. Cocking, S. Flessa, G. Reinelt, Locating health facilities in Nouna district, Burkina Faso, in: Operations Research Proceedings, 2005, Part X, 2006, pp. 431–436.
[40] A. Mehretu, R. Wittick, B. Pigozzi, Spatial design for basic needs in eastern upper volta, The Journal of Developing Areas 17 (1983) 383–394.
[41] J. Oppong, Accommodating the rainy season in third world location-allocation applications, Socio-Economic Planning Sciences 30 (1996) 121–137.
[42] S. Rahman, D. Smith, Use of location-allocation models in health service development planning in developing nations, European Journal of Operational Research 123 (2000) 437–452.
[43] G. Rushton, Use of location-allocation models for improving the geographical accessibility of rural services in developing countries, International Regional Science Review 9 (1984) 217–240.
[44] J. Tien, K. El-Tell, A quasi hierarchical location-allocation model for primary health care planning, IEEE Transactions on Systems, Man, and Cybernetics 14 (1984) 373–380.
[45] Statistical Center of Iran, Census report of Iran population. http://www.amar.org.ir, 2007.