Computers and Electrical Engineering 34 (2008) 470–487

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

A tabu search algorithm for the global planning problem of third generation mobile networks

Marc St-Hilaire [a], Steven Chamberland [b,*], Samuel Pierre [b]

[a] Department of Systems and Computer Engineering, Carleton University, 1125 Colonel By Drive, Ottawa, Ontario, Canada K1S 5B6
[b] Department of Computer Engineering, École Polytechnique de Montréal, C.P. 6079, Succ. Centre-Ville, Montréal, Québec, Canada H3C 3A7

ARTICLE INFO

Article history:
Received 24 August 2006
Accepted 2 February 2008
Available online 22 April 2008

Keywords:
Third generation (3G) mobile networks
Universal mobile telecommunications system (UMTS)
Network planning
Global approach
Metaheuristic
Tabu search (TS)

ABSTRACT

In this paper, we propose a tabu search (TS) algorithm for the global planning problem of third generation (3G) universal mobile telecommunications system (UMTS) networks. This problem is composed of three NP-hard subproblems: the cell, the access network and the core network planning subproblems. Therefore, the global planning problem consists in selecting the number, the location and the type of network nodes (including the base stations, the radio network controllers, the mobile switching centers and the serving GPRS (General Packet Radio Service) support nodes) as well as the interconnections between them. After describing our metaheuristic, a systematic set of experiments is designed to assess its performance. The results show that quasi-optimal solutions can be obtained with the proposed approach.

1. Introduction

Before deploying or updating any cellular network infrastructures, it is crucial to carefully plan this step. Maximizing the coverage, providing a good quality of service and minimizing the cost are just a few reasons to justify an adequate planning. As a result, planning tools can be very useful for large and complex infrastructures such as the third generation (3G) universal mobile telecommunications system (UMTS) networks.

As shown in Fig. 1, UMTS networks are typically composed of two different parts: the access network and the core network. Also called universal terrestrial radio access network (UTRAN), the access network is composed of base stations (also called node Bs) and radio network controllers (RNCs). The node Bs are mainly used to transmit/receive radio frequencies to/from the mobile users (MUs). More specifically, they deal with channel coding, rate adaptation, spreading, etc. The wideband code division multiple access (WCDMA) scheme is used as the air-interface between the users and the base stations. The latter is using the direct sequence CDMA (DS-CDMA) to provide higher speed (up to 2 Mb/s) and support more users compared to previous network generations. In WCDMA, the information sent is spread over a wideband of around 5 MHz. The RNCs deal with resource and mobility management (load balancing, admission control, code allocation, etc.) as well as concentrating and forwarding the traffic to the core network.

As depicted in Fig. 1, the core network is also subdivided in two different parts. These two parts inherit from previous network generations (see Table 1). In fact, the circuit switched network is the heritage of the global system for mobile communications (GSM) architecture while the packet switched network is the legacy of the general packet radio system (GPRS) infrastructure. In terms of equipment and utility, the circuit switched part is composed of mobile switching centers (MSCs)

* Corresponding author.
E-mail addresses: marc_st_hilaire@carleton.ca (M. St-Hilaire), steven.chamberland@polymtl.ca (S. Chamberland), samuel.pierre@polymtl.ca (S. Pierre).

0045-7906/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compeleceng.2008.02.001

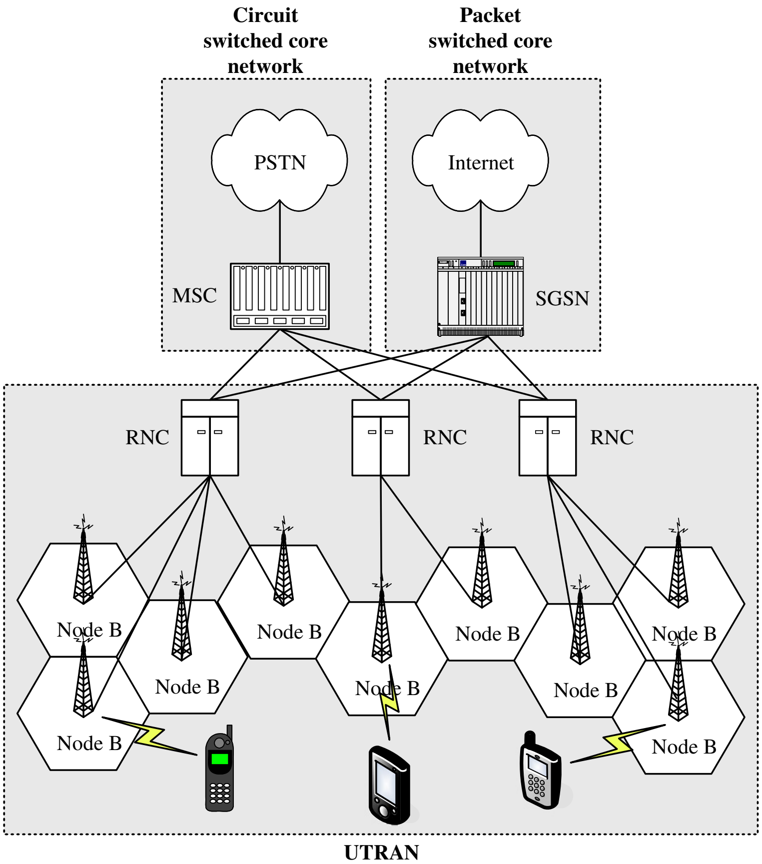

Fig. 1. The UMTS network architecture.

Table 1
Cellular network evolution

	2G	2.5G	3G
Main standard	GSM	GPRS	UMTS
Access technology	TDMA	TDMA	WCDMA
Speed (kb/s)	9.6	171.2	2000
Type of traffic	Voice	Voice and data	Voice and data

and provide access to the public switched telephone network (PSTN) while the packet switched part is composed of serving GPRS support nodes (SGSNs) and provides access to the Internet.

The UMTS network planning problem has been widely studied in the literature. In fact, several models and exact/approximate algorithms have been proposed to solve these problems. However, the majority of these studies only focus on a portion of the overall problem. Therefore, it is common to find papers about the different subproblems such as the cell, the access network and the core network planning subproblems. The following describe the technological aspects and the main challenges that need to be considered when planning each subproblem:

- *The cell planning subproblem*: When planning the radio part, the most important task is to consider the coverage and the capacity. In order to do that, we need to do an analysis of the radio links. This include the use of a propagation model and different link budget parameters. The goal is to find the optimal location of the sites, the antenna height, the tilt, the orientation, etc. Another important task is the radio resource management. This involves the admission control, the power control and handovers. In order to carefully plan these tasks, it is important to have an accurate prediction of the number of users. The most interesting papers about this subproblems are Amaldi et al. [1,2] and Thiel et al. [16].
- *The access network planning subproblem*: In general, when planning the access network, we first need to configure the network elements (RNC). This include, but not limited to, the location, the number and the capacity of equipment needed. Then, we need to consider the notion of homing (single or double) which state how the base stations will be linked to

the RNC. Proper attention to the different types of traffic is also important and even different classes of services can be used for the data. Finally, traffic management and parameters must be considered in order to make sure that every client will receive a proper quality of service. The most interesting papers about this subproblems are Harmatos et al. [10], Lauther et al. [14] and Wu and Pierre [17,18].

- *The core network planning subproblem*: This problem is similar to fixed network planning problems. It involves dimensioning the network elements (MSC and SGSN) and the interconnection between them. Different link types (optical, ethernet) with different capacities can be used. Before sizing the links, we first need to know where the traffic will be sent (i.e., which routing strategy will be used). The most interesting paper about this subproblem is Harmatos [9].

For an extensive review of the literature, see St-Hilaire et al. [15].

Actually, the global planning approach has been considered only by St-Hilaire et al. [15]. The authors have proposed a mathematical model for the global problem and they have shown that only small-size instances of the problem can be solved to optimality within a reasonable amount of time using a branch-and-bound algorithm. Moreover, the authors have also proposed a local search (LS) algorithm to find feasible solutions rapidly. The numerical results have demonstrated that this algorithm can find solutions that are, on average, at 6.53% of the optimal solution.

The aim of this paper is to further improve the results obtained with the LS algorithm as proposed in [15]. To achieve this goal, we explore the concept of the tabu search (TS) algorithm and apply it to the global planning problem of UMTS networks. The main idea behind this algorithm is to find an acceptable tradeoff between the quality of the solution (i.e., the closeness to the optimum) and the time required to find an answer. The proposed TS algorithm deals simultaneously with the three subproblems mentioned previously in order to find "good" feasible solutions rapidly.

The rest of this paper is organized as follows. In Section 2, we present a brief overview of the cell, the access network and the core network planning subproblems. In Section 3, we propose and present a metaheuristic algorithm based on the tabu search principle. In Section 4, computational results are presented and analyzed, and finally, conclusions and further works are presented in Section 5.

2. Network planning subproblems

In this section, we present a brief overview of the three subproblems contained in the UMTS network planning problem. It is important to note that each of these subproblems is NP-hard (see [1,9,10]).

The uplink direction of the traffic (i.e., from the MUs to node Bs) is considered in this paper. The uplink traffic is very important when the amount of data exchanged is balanced between the uplink and the downlink directions. To simulate the traffic, we introduce the notion of test points (TPs). Each TP can be viewed as a centroid where a given amount of traffic is requested. Therefore, one TP can represent several MUs in a given area. When planning the uplink direction, the restrictions are the MU transmit power and the interference (see [11]).

Before presenting each subproblems, we need to define a notation. This notation will be used to represent the objective function of each subproblem. The notation is composed of sets, decision variables and cost parameters (see Appendix A for the complete notation).

- Sets:
  - S_1,S_2,S_3,S_4, respectively the set of potential sites to install the node Bs, the RNCs, the MSCs and the SGSNs. Most of the time, service providers already have a predefined set of possible locations. For example, they might install a base station on the top of a building they already own;
  - T_1,T_2,T_3,T_4, respectively the set of node B types, RNC types, MSC types and SGSN types. This allows different equipment (in terms of cost, power, sensitivity and capacity) to be used;
  - M_{12},M_{23},M_{24}, respectively the set of links and interface types that can be used to connect the node Bs to the RNCs, the RNCs to the MSCs and the RNCs to the SGSNs. Different links (with different capacities) can be used in the planning process in order to meet the traffic requirements.
- Decision variables:
  - v_{12}^{ijm}, the number of links of type $m \in M_{12}$ connecting the node B installed at site $i \in S_1$ to the RNC installed at site $j \in S_2$;
  - v_{23}^{jkm}, the number of links of type $m \in M_{23}$ connecting the RNC installed at site $j \in S_2$ to the MSC installed at site $k \in S_3$;
  - v_{24}^{jlm}, the number of links of type $m \in M_{24}$ connecting the RNC installed at site $j \in S_2$ to the SGSN installed at site $l \in S_4$;
  - x_1^{it}, a 0–1 variable such that $x_1^{it} = 1$ if and only if a node B of type $t \in T_1$ is installed at site $i \in S_1$. Basically, this variable (and the three below) indicate if a node B (RNC, MSC, SGSN) is installed at a given location;
  - x_2^{jt}, a 0–1 variable such that $x_2^{jt} = 1$ if and only if an RNC of type $t \in T_2$ is installed at RNC site $j \in S_2$;
  - x_3^{kt}, a 0–1 variable such that $x_3^{kt} = 1$ if and only if an MSC of type $t \in T_3$ is installed at MSC site $k \in S_3$;
  - x_4^{lt} be a 0–1 variable such that $x_4^{lt} = 1$ if and only if an SGSN of type $t \in T_4$ is installed at site $l \in S_4$.

- Cost parameters:
  - a_{12}^{ijm}, the link and interface costs (including installation cost) for connecting a node B installed at site $i \in S_1$ to an RNC installed at site $j \in S_2$ through a link and interface of type $m \in M_{12}$. The interface cost is a fixed cost while the link cost is proportional to the distance between the two locations;
  - a_{23}^{jkm}, the link and interface costs (including installation cost) for connecting an RNC installed at site $j \in S_2$ to an MSC installed at site $k \in S_3$ through a link and interface of type $m \in M_{23}$;
  - a_{24}^{jlm}, the link and interface costs (including installation cost) for connecting an RNC installed at site $j \in S_2$ to an SGSN installed at site $l \in S_4$ through a link and interface of type $m \in M_{24}$;
  - $b_1^t, b_2^t, b_3^t, b_4^t$, respectively the cost (including installation cost) of a node B of type $t \in T_1$, of an RNC of type $t \in T_2$, of an MSC of type $t \in T_3$ and of an SGSN of type $t \in T_4$.

2.1. The cell planning subproblem

In order for the users to communicate with the base stations, an access mechanism to the medium is required. In UMTS networks, the WCDMA scheme is typically used. With this scheme, the capacity of each cell is based on the interference levels (for more information, see Amaldi et al. [1]). Different types of antenna can be used. Typical antennas are either directional or omnidirectional depending on the coverage needed.

Given a set of potential node B locations, the cell planning problem consists in finding the number, the location and the type of node Bs subject to signal quality and coverage constraints. The objective is to minimize the cost of node Bs (including the installation cost). The cell planning subproblem can be formulated as follows:

Objective:

$$\min \sum_{t \in T_1} b_1^t \sum_{i \in S_1} x_1^{it} \tag{1}$$

Subject to the following assumptions and constraints (see Appendix A for the mathematical formulation of these constraints):

- each TP is associated to exactly one node B. This means that each test point (mobile user) has to be assigned to one base station. In fact, no TPs can be left uncovered. This assignment is also called single homing (as opposed to multi homing where mobile users can be assigned to more than one base station);
- at most one node B can be installed at a node B site. In this paper, only a single equipment (a single node B) can be installed at any given physical location. This assumption could easily be replaced in order to allow multiple equipment to be installed at the same physical location;
- the sum of the traffic (in circuits and in b/s) from the TPs associated to a node B cannot exceed the capacity (in circuits and in b/s) of that node B. The base stations have a finite capacity and we simply want to make sure that the total traffic can be handled by the base station;
- signal power and signal to interference ratio (SIR) constraints. Since we are considering the uplink direction, the emitted power of the mobile user minus the propagation loss should be high enough so that the base station can receive the signal with enough power. Also, intercell and intracell interferences should be kept as low as possible.

When considering the cell planning subproblem, we also suppose the following information is known (usually given as input):

- the location of the TPs;
- the estimated traffic from each TP for each class of traffic;
- the set of possible sites to install the node Bs;
- the different types of node Bs;
- the cost of the node B types (including, for instance, floor space, cables, racks, patch panels, electrical installations, labor, etc.).

2.2. The access network planning subproblem

The access network planning subproblem consists in determining:

- the number, the location and the types of the RNCs;
- the assignment of node Bs to the RNCs;
- the number, the location and the type of links used to connect node Bs to the RNCs.

The objective is to minimize the cost of the access network (including the cost of the RNCs and the cost of the links and interfaces). This subproblem is similar to the cell to switch assignment problem in 2G networks. The model for the access network planning problem is the following:

Objective:

$$\min\left(\sum_{t\in T_2} b_2^t \sum_{j\in S_2} x_2^{jt} + \sum_{i\in S_1}\sum_{j\in S_2}\sum_{m\in M_{12}} a_{12}^{ijm} v_{12}^{ijm}\right) \tag{2}$$

Subject to the following assumptions and constraints (see Appendix A for the mathematical formulation of these constraints):

- each node B is connected to exactly one RNC. We made this assumption in order to simplify the model. This could easily be replaced by a constraint where each node B is linked to two RNCs in order to add reliability to the network;
- the number of links connected to a node B cannot exceed the maximum number of interfaces that can be installed in that node B. Since the model is using modular equipment (i.e., different interface cards can be installed), the number of interface cannot exceed the number of available slots;
- at most one RNC can be installed at an RNC site;
- due to equipment modularity, the number of links connected to an RNC cannot exceed the maximum number of interfaces that can be installed in that RNC;
- the sum of the capacities of the links (in circuits and in b/s) connected to an RNC cannot exceed its capacity (in circuits and in b/s). In other words, each network equipment has a maximal amount of traffic that it can handle at the same time (called the switch fabric capacity);
- the traffic from a node Bs connected to a particular RNC cannot exceed the capacity of the link between these two locations.

When planning the access network, we suppose the following information is known (i.e., given as input):

- the location of node Bs;
- the traffic from each node B;
- the set of possible sites to install the RNCs;
- the different types of RNCs;
- the cost of the RNC types (including floor space, cables, racks, patch panels, electrical installations, labor, etc.).

2.3. The core network planning subproblem

The core network planning subproblem consists in determining:

- the number, the location and the type of the MSCs and SGSNs;
- the assignment of the RNCs to the MSCs and SGSNs;
- the number, the location and the type of links used to connect the RNCs to the MSCs and SGSNs.

The objective is to minimize the cost of the core network (including the cost of the MSCs, the cost of the SGSNs and the cost of the links and interfaces). The model for the core network planning problem is the following:

Objective:

$$\min\left(\sum_{t\in T_3} b_3^t \sum_{k\in S_3} x_3^{kt} + \sum_{t\in T_4} b_4^t \sum_{l\in S_4} x_4^{lt} + \sum_{j\in S_2}\sum_{k\in S_3}\sum_{m\in M_{23}} a_{23}^{jkm} v_{23}^{jkm} + \sum_{j\in S_2}\sum_{l\in S_4}\sum_{m\in M_{24}} a_{24}^{jlm} v_{24}^{jlm}\right) \tag{3}$$

Subject to the following assumptions and constraints (see Appendix A for the mathematical formulation of these constraints):

- each RNC is connected to exactly one MSC and one SGSN. As stated before, this assumption is used to simplify the model. It could easily be replaced by any other connectivity constraints;
- each MSC (SGSN) is connected to the PSTN (PDN);
- at most one MSC (SGSN) can be installed at an MSC (SGSN) site;
- Due to equipment modularity, the number of links connected to an MSC (SGSN) cannot exceed the maximum number of interfaces that can be installed in that MSC (SGSN);
- the sum of the capacities of the links (in circuits) connected to an MSC cannot exceed its maximal capacity (in circuits);
- the sum of the capacities of the links (in b/s) connected to an SGSN cannot exceed its maximal capacity (in b/s).

When planning the core network, we suppose the following information is known (i.e., given as input):

- the location of the RNCs;
- the traffic from each RNC;
- the set of possible sites to install the MSCs and SGSNs;
- the different types of MSCs and SGSNs;
- the cost of the MSC and SGSN types (including floor space, cables, racks, patch panels, electrical installations, labor, etc.).

When solving each subproblem individually, we need to use all the information mentioned above. However, the information required by the global planning problem is not the sum of all information required for each individual subproblem. In fact, when the three subproblems are connected under a unique global formulation (i.e., a global approach), only the following information is required as input:

- the location of the TP as well as the requested amount of traffic per TP;
- the possible locations to install the network equipment (node B, RNC, MSC and SGSN);
- the different types of equipment that can be used in the planning process (node B, RNC, MSC, SGSN, interface cards and links);
- the cost of the network equipment (node B, RNC, MSC, SGSN, interface cards and links).

In a global approach, the cell, the access and the core network planning problems are solved simultaneously and consequently, the different interactions between them are taken into consideration. The mathematical model for the global planning problem of UMTS networks (denoted GPU) has been recently proposed by St-Hilaire et al. [15] (see also Appendix A). Since this problem is NP-hard, it is unlikely that real-size instances of the problem can be solved to optimality within a reasonable amount of time [15]. As a result, in the next section, we propose a TS algorithm that deals simultaneously with these subproblems in order to find "good" feasible solutions.

3. The tabu search algorithm

In this section, we propose a TS algorithm for the global planning problem of UMTS networks. For an introduction to TS, see Glover [7,8]. The simple TS algorithm uses the best improvement local search as basic ingredient and uses a short term memory to escape from local minima [3]. Starting from the current solution, TS finds a better one in its neighborhood. A neighborhood is a set of solutions that are found by applying an appropriate transformation of the current solution. In order for the algorithm to move away from a local minimum, the search allows moves resulting in a degradation of the objective function value, thus avoiding the trap of local optimality. To prevent the search from cycling, solutions obtained recently and moves that reverse the effect of recent moves are considered "tabu".

In the next subsection, we propose a decomposition approach to solve GPU($\mathbf{x}$), i.e., the model GPU when the vector $\mathbf{x} = (\mathbf{x_1}, \mathbf{x_2}, \mathbf{x_3}, \mathbf{x_4})$ is fixed (see Appendix A for the notation). The vector $\mathbf{x}$ is used to represent the actual topology of the network. Given a geographical area with possible locations to install equipment (as shown in Fig. 2), the vector $\mathbf{x}$ indicates if an equipment is installed at a given location. As we can see in Fig. 2, four possible node B locations are available but only two are used (shaded ones). This results in the first four number of vector $\mathbf{x}$. The same process applies to the RNC, MSC and SGSN until all locations have been processed.

3.1. Solving GPU($\mathbf{x}$)

When the vector $\mathbf{x}$ is fixed (i.e., the locations and the types of the nodes are fixed), GPU can be decomposed into four subproblems: the TPs to node Bs assignment subproblem; the node Bs to RNCs assignment subproblem; the RNCs to MSCs assign-

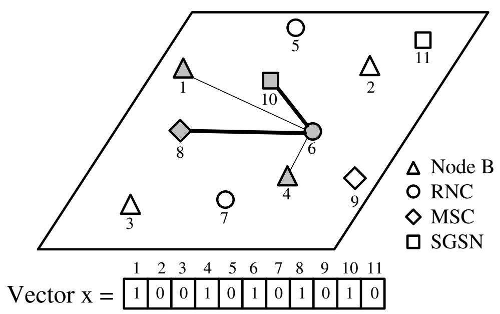

Fig. 2. Representation of the vector **x**.

ment subproblem and the RNCs to SGSNs assignment subproblem. The first subproblem consists of assigning the TPs to node Bs while satisfying the power, the capacity and the SIR constraints. This subproblem is NP-hard (transformation from the knapsack problem) and to solve it, we use the CPLEX Mixed Integer Optimizer 9.0 [12]. The other assignment subproblems are also NP-hard (see [4,5]) and to solve them, we use the shortest augmenting path algorithm LAPJV of Jonker and Volgenant [13].

3.2. The tabu search algorithm

Let $s \in S$ be a solution. $N(s)$ is called the neighborhood of s and consists of the solutions obtained by performing all possible moves. Each move consists in modifying one variable of the vector $\mathbf{x}$. The following types of move are considered:

- remove a node (a node B, an RNC, an MSC or an SGSN that is already installed);
- add a node (a node B, an RNC, an MSC or an SGSN of type t);
- change the type of a node (a node B, an RNC, an MSC or an SGSN that is already installed);
- change the type of a node (a node B, an RNC, an MSC or an SGSN that is already installed);
- swap a node which is already installed at a given location to a vacant location.

Table 2
Features of the node B types

Characteristics	Type A	Type B	Type C
Capacity (circuits)	100	200	400
Capacity (Mb/s)	120	240	480
Number of interfaces	1	2	2
Sensitivity (dBm)	−90	−100	−110
Cost ($)	20,000	30,000	50,000

Table 3
Features of the RNC types

Characteristics	Type A	Type B	Type C
Switch fabric capacity (Mb/s)	2000	5000	10,000
Number of node B interfaces	10	20	40
Number of MSC/SGSN interfaces	15	30	60
Cost ($)	50,000	90,000	120,000

Table 4
Features of the MSC types

Characteristics	Type A	Type B	Type C
Switch fabric capacity (circuits)	100,000	200,000	300,000
Number of interfaces	50	100	150
Cost ($)	200,000	350,000	500,000

Table 5
Features of the SGSN types

Characteristics	Type A	Type B
Switch fabric capacity (Mb/s)	20,000	40,000
Number of interfaces	16	32
Cost ($)	40,000	60,000

Table 6
Cost of the links

Link type	Capacity	Cost ($/km)
DS-3	2688 circuits	2500
OC-3	155 Mb/s	1500
OC-12	622 Mb/s	4000
GE	1 Gb/s	4000

At each iteration, we select the best move (among all possible moves) while taking into consideration the tabus and the aspiration criteria. The chosen site is declared tabu for a number of iterations that is randomly determined according to a uniform discrete distribution specified on an interval. Therefore, over multiple runs, TS may find different solutions. As a result, a multi-start TS is proposed.

Table 7
Cost of the interface types

Interface type	Cost ($)
DS-3	1500
OC-3	2000
OC-12	4500
GE	2000

Table 8
Problem sizes

Problem #	Number of TPs	Number of node Bs	Number of RNCs	Number of MSCs	Number of SGSNs
1	5	10	5	5	5
2	5	10	10	10	10
3	5	20	5	5	5
4	5	20	10	10	10
5	5	30	5	5	5
6	5	30	10	10	10
7	10	10	5	5	5
8	10	10	10	10	10
9	10	20	5	5	5
10	10	20	10	10	10
11	10	30	5	5	5
12	10	30	10	10	10
13	15	10	5	5	5
14	15	10	10	10	10
15	15	20	5	5	5
16	15	20	10	10	10
17	15	30	5	5	5
18	15	30	10	10	10
19	20	10	5	5	5
20	20	10	10	10	10
21	20	20	5	5	5
22	20	20	10	10	10
23	20	30	5	5	5
24	20	30	10	10	10
25	25	10	5	5	5
26	25	10	10	10	10
27	25	20	5	5	5
28	25	20	10	10	10
29	25	30	5	5	5
30	25	30	10	10	10
31	30	10	5	5	5
32	30	10	10	10	10
33	30	20	5	5	5
34	30	20	10	10	10
35	30	30	5	5	5
36	30	30	10	10	10
37	35	10	5	5	5
38	35	10	10	10	10
39	35	20	5	5	5
40	35	20	10	10	10
41	35	30	5	5	5
42	35	30	10	10	10
43	40	10	5	5	5
44	40	10	10	10	10
45	40	20	5	5	5
46	40	20	10	10	10
47	40	30	5	5	5
48	40	30	10	10	10

The aspiration criteria states that if the use of a tabu site allow us to find a better solution than any other found so far, we remove the tabu on this site, since there is no risk of cycling. We will now proceed to the detailed description of the algorithm:

TS Algorithm

Step 1: (Initial solution)Find an initial solution using the LS algorithm proposed in [15].

Repeat Steps 2–3 for Number_Starts iterations

Step 2: (TS)

Repeat Steps 2.1–2.2 for Number_Iterations iterations

Step 2.1: (Exploring the neighborhood)

2.1.1 Determine the best move, while taking into consideration the tabu moves and the aspiration criteria. For each move $\mathbf{x} \rightarrow \mathbf{x}'$, we find the solution by solving GPU($\mathbf{x}'$). The cost of the solution is the total cost of the network.

2.1.2 Determine the number of iterations (according to a uniform distribution) for which the chosen site is tabu.

Table 9
Computational results (test #1 and test #2)

Problem #	Test #1					Test #2				
	CPLEX		TS		GAP (%)	CPLEX		TS		GAP (%)
	Value ($)	CPU (s)	Value ($)	CPU (s)		Value ($)	CPU (s)	Value ($)	CPU (s)	
1	354,045	3	354,045	59	0.00	354,996	1	354,996	46	0.00
2	407,640	176	407,640	151	0.00	408,121	253	408,121	119	0.00
3	353,922	24	353,922	130	0.00	353,152	14	353,152	101	0.00
4	352,807	275	353,213	230	0.12	378,373	333	378,624	188	0.07
5	355,880	37	355,880	232	0.00	353,454	22	353,454	200	0.00
6	377,601	576	377,601	396	0.00	378,047	467	378,753	331	0.19
7	435,908	9	435,908	73	0.00	380,093	11	380,093	52	0.00
8	378,325	58	378,325	154	0.00	388,251	104	388,286	119	0.01
9	410,286	29	410,286	207	0.00	408,793	67	408,793	165	0.00
10	352,890	466	353,231	329	0.10	378,933	267	378,933	258	0.00
11	389,682	66	391,140	501	0.37	379,948	49	379,948	367	0.00
12	352,420	242	352,420	646	0.00	402,306	3228	402,306	467	0.00
13	408,705	4	408,705	78	0.00	408,913	8	408,913	68	0.00
14	443,421	611	443,421	171	0.00	433,651	1049	433,651	137	0.00
15	408,274	61	408,274	252	0.00	408,675	95	424,220	230	3.80
16	406,914	4173	406,914	466	0.00	388,124	1584	389,160	298	0.27
17	407,005	149	407,005	732	0.00	389,752	96	389,752	470	0.00
18	408,448	10,314	408,448	928	0.00	403,091	11,677	407,200	712	1.02
19	433,388	72	433,388	102	0.00	443,620	12	443,620	79	0.00
20	443,900	463	443,900	207	0.00	439,654	613	439,654	146	0.00
21	434,191	256	434,191	352	0.00	424,762	376	424,762	283	0.00
22	434,800	2692	443,767	505	2.06	TL	108,000	443,813	394	–
23	426,377	313	443,511	1043	4.02	425,271	3075	425,271	723	0.00
24	TL	108,000	443,841	1248	–	432,809	16,494	433,051	905	0.06
25	499,261	28	499,261	123	0.00	469,239	46	469,239	86	0.00
26	444,452	107	444,825	224	0.08	459,955	3396	460,370	181	0.09
27	467,535	507	469,471	496	0.41	465,521	233	479,639	356	3.03
28	443,135	23,544	443,135	607	0.00	504,768	8832	525,256	476	4.06
29	449,794	20,653	449,794	1301	0.00	445,092	13,326	446,612	847	0.34
30	449,020	20,028	470,375	1475	4.76	469,358	19,324	469,907	1216	0.12
31	505,226	856	505,226	147	0.00	494,204	713	494,204	122	0.00
32	514,312	8822	515,016	259	0.14	514,245	11,665	514,245	189	0.00
33	481,116	6385	481,151	587	0.01	494,761	26,235	495,804	433	0.21
34	484,236	25,826	502,809	813	3.84	519,002	15,191	528,765	618	1.88
35	TL	108,000	461,927	1547	–	480,259	72,722	480,259	1207	0.00
36	TL	108,000	477,908	1777	–	TL	108,000	479,500	1220	–
37	514,474	74	514,474	152	0.00	615,059	152	615,059	117	0.00
38	513,898	1858	513,898	259	0.00	TL	108,000	558,353	207	–
39	514,925	42,731	519,875	776	0.96	505,605	1023	505,605	449	0.00
40	TL	108,000	694,478	1004	–	TL	108,000	534,629	715	–
41	484,759	62,458	499,805	2141	3.10	TL	108,000	520,123	1576	–
42	TL	108,000	554,877	1739	–	TL	108,000	503,741	1750	–
43	555,559	3	555,559	174	0.00	590,471	394	590,471	125	0.00
44	520,429	31,616	520,888	290	0.09	569,623	694	569,623	228	0.00
45	522,172	81,681	522,172	807	0.00	514,512	4004	514,512	782	0.00
46	TL	108,000	529,581	1179	–	TL	108,000	540,834	773	–
47	TL	108,000	558,363	1777	–	TL	108,000	542,707	1587	–
48	TL	108,000	594,314	2077	–	TL	108,000	538,531	2127	–

Step 2.2: (TS best solution update)If the cost of the current solution is less than the cost of the best solution found so far, update this best solution.

Step 3: (Multi-start best solution update)

If the cost of the current solution is less than the cost of the best solution found so far, update this best solution.

4. Computational results

In this section, we present the results of a systematic set of experiments designed to assess the performance of the TS algorithm.

For the tests, the signal propagation model proposed in [6] is used. Moreover, three node B types, three RNC types, three MSC type and two SGSN types are available. Their features are respectively presented in Tables 2–5. Moreover, OC-3 and

Table 10
Computational results (test #3 and test #4)

Problem #	Test #3					Test #4				
	CPLEX		TS		GAP (%)	CPLEX		TS		GAP (%)
	Value ($)	CPU (s)	Value ($)	CPU (s)		Value ($)	CPU (s)	Value ($)	CPU (s)	
1	352,390	1	352,390	39	0.00	353,948	4	353,948	40	0.00
2	388,510	169	388,510	99	0.00	352,865	23	353,453	115	0.17
3	353,295	11	353,295	79	0.00	354,319	30	354,319	84	0.00
4	353,208	1256	353,208	153	0.00	388,283	323	388,283	188	0.00
5	353,099	11	353,099	169	0.00	390,489	96	408,935	207	4.72
6	377,743	712	377,743	257	0.00	377,826	1563	377,826	328	0.00
7	435,058	17	435,058	46	0.00	408,082	3	408,082	87	0.00
8	409,266	207	434,019	103	6.05	408,198	103	408,198	118	0.00
9	379,464	31	379,464	151	0.00	380,443	46	381,212	119	0.20
10	388,406	2785	388,406	225	0.00	378,481	814	378,481	236	0.00
11	390,052	35	390,052	328	0.00	379,446	78	379,446	303	0.00
12	378,497	3001	378,497	459	0.00	404,012	3230	404,012	529	0.00
13	408,783	7	408,783	62	0.00	443,097	16	443,097	71	0.00
14	433,270	184	433,270	122	0.00	444,748	272	464,132	137	4.36
15	407,776	64	407,776	190	0.00	408,372	596	413,590	204	1.28
16	408,498	1040	408,534	311	0.01	433,178	632	433,178	359	0.00
17	388,437	44,091	388,437	456	0.00	409,390	191	442,358	492	8.05
18	387,794	44,023	387,794	661	0.00	408,530	3702	433,295	619	6.06
19	424,602	115	424,605	73	0.00	425,277	65	434,500	58	2.17
20	459,237	2490	462,343	143	0.68	423,112	1399	423,112	151	0.00
21	410,257	165	410,257	227	0.00	415,847	3044	435,308	247	4.68
22	TL	108,000	429,272	402	–	438,518	1298	443,420	408	1.12
23	445,202	4813	445,202	697	0.00	424,344	16,362	424,344	627	0.00
24	TL	108,000	443,640	729	–	423,574	20,990	423,574	866	0.00
25	490,882	25	490,882	91	0.00	494,692	51	494,692	78	0.00
26	484,489	1392	484,489	144	0.00	462,704	2199	462,704	177	0.00
27	460,753	2336	460,753	300	0.00	480,743	889	480,743	362	0.00
28	443,040	22,156	443,051	422	0.00	TL	108,000	498,769	505	–
29	444,703	1035	460,071	873	3.46	469,355	1014	483,839	1013	3.09
30	TL	108,000	474,475	1096	–	TL	108,000	448,681	1139	–
31	498,555	198	498,555	112	0.00	468,999	33	480,125	99	2.37
32	504,282	12,320	504,282	227	0.00	494,165	2959	494,304	189	0.03
33	473,435	2464	488,190	410	3.12	467,888	11,274	467,888	443	0.00
34	TL	108,000	510,337	494	–	TL	108,000	538,388	550	–
35	479,497	66,151	484,947	1217	1.14	460,560	6,558	460,560	1167	0.00
36	TL	108,000	494,663	1316	–	TL	108,000	492,175	1548	–
37	576,884	4	576,884	107	0.00	507,799	460	515,137	105	1.45
38	TL	108,000	568,679	187	–	550,053	2625	558,669	211	1.57
39	496,602	11,710	496,602	407	0.00	TL	108,000	539,404	566	–
40	TL	108,000	519,217	672	–	TL	108,000	528,763	688	–
41	TL	108,000	510,324	1331	–	TL	108,000	540,005	1515	–
42	TL	108,000	519,102	1578	–	TL	108,000	494,975	1807	–
43	594,929	1343	601,188	112	1.05	568,630	59	568,630	108	0.00
44	493,344	501	493,344	294	0.00	548,202	2,393	553,378	226	0.94
45	TL	108,000	540,722	538	–	526,654	5,693	527,821	683	0.22
46	TL	108,000	564,061	834	–	TL	108,000	508,285	787	–
47	TL	108,000	526,029	1777	–	TL	108,000	523,067	1822	–
48	TL	108,000	550,097	2000	–	TL	108,000	500,480	2013	–

OC-12 links can be used to connect node Bs to the RNC, DS-3 links are used to connect the RNCs to the MSC and gigabit Ethernet (GE) links are used to connect the RNCs to the SGSNs (see Table 6). The cost of the types of interfaces are presented in Table 7.

As shown in Table 8, 48 different problem sizes were selected for the tests. For each size, five instances of the problem were randomly generated on a 4 km^2 region. Each instance is solved with two different methods. For the first one, we used the CPLEX Mixed Integer Optimizer 9.0 (see [12] for more information about CPLEX) to solve the model GPU presented in Appendix A. The algorithm used by CPLEX is the branch-and-bound algorithm. This algorithm gives us the optimal solution and therefore, a reference point to evaluate the solution found using the TS algorithm. Since CPLEX can be computationally intensive, the CPU time limit of the branch-and-bound algorithm is set to 30 h. It is important to note that the solutions returned after the time limit might not be optimal. The second method used is the TS algorithm. The Number_Starts parameter of the TS algorithm is set to 3, the Number_Iterations parameter to 100 and the interval of the number of iterations that a site can be tabu is set to [5,9]. These parameters were selected after several experimentations. In fact, the best solutions were found with those parameters. All tests are carried out on a linux workstation with a 3 GHz CPU and 1 GB of RAM.

The results are presented in Tables 9–11. The first column presents the problem number. The next two columns present, respectively, the optimal solution found by CPLEX (if the time limit was not reached) and the CPU execution time. The fol-

Table 11
Computational results (test #5 and average gap)

Problem #	Test #5					GAP		
	CPLEX		TS		GAP (%)	GAP Min (%)	GAP Max (%)	GAP Avg (%)
	Value ($)	CPU (s)	Value ($)	CPU (s)				
1	353,665	6	353,665	41	0.00	0.00	0.00	0.00
2	352,708	78	353,508	95	0.23	0.00	0.23	0.08
3	388,656	31	388,656	99	0.00	0.00	0.00	0.00
4	353,241	91	353,241	155	0.00	0.00	0.12	0.04
5	353,806	33	353,806	206	0.00	0.00	4.72	0.94
6	352,587	248	352,587	248	0.00	0.00	0.19	0.04
7	381,004	20	381,004	53	0.00	0.00	0.00	0.00
8	387,879	311	387,879	109	0.00	0.00	6.05	1.21
9	379,149	19	379,182	150	0.01	0.00	0.20	0.04
10	352,311	120	352,361	224	0.01	0.00	0.10	0.02
11	405,167	73	405,167	291	0.00	0.00	0.37	0.07
12	378,328	1129	378,328	435	0.00	0.00	0.00	0.00
13	412,103	69	412,103	59	0.00	0.00	0.00	0.00
14	407,206	39	407,206	199	0.00	0.00	4.36	0.87
15	408,501	235	409,603	208	0.27	0.00	3.80	1.07
16	389,031	1079	389,044	273	0.00	0.00	0.27	0.06
17	410,038	916	416,404	474	1.55	0.00	8.05	1.92
18	403,487	4,736	408,096	608	1.14	0.00	6.06	1.64
19	424,534	99	424,534	69	0.00	0.00	2.17	0.43
20	443,490	333	443,490	136	0.00	0.00	0.68	0.14
21	447,015	603	448,082	263	0.24	0.00	4.68	0.98
22	443,347	2,253	444,924	406	0.36	0.36	2.06	1.18
23	422,913	6,443	422,913	664	0.00	0.00	4.02	0.80
24	424,203	102,067	424,203	909	0.00	0.00	0.06	0.02
25	469,877	77	469,877	80	0.00	0.00	0.00	0.00
26	515,012	3492	515,012	149	0.00	0.00	0.09	0.03
27	TL	108,000	480,238	369	-	0.00	3.03	0.86
28	443,463	4873	443,463	478	0.00	0.00	4.06	1.02
29	TL	108,000	481,002	831	-	0.00	3.46	1.72
30	459,583	12,093	459,583	988	0.00	0.00	4.76	1.63
31	515,731	6	515,731	79	0.00	0.00	2.37	0.47
32	553,646	34,723	553,646	176	0.00	0.00	0.14	0.03
33	459,434	3131	459,434	376	0.00	0.00	3.12	0.67
34	488,991	68,602	504,701	491	3.21	1.88	3.84	2.98
35	480,028	50,723	480,271	1207	0.05	0.00	1.14	0.30
36	TL	108,000	460,849	1237	–	–	–	–
37	532,295	156	532,295	113	0.00	0.00	1.45	0.29
38	584,564	2962	584,564	196	0.00	0.00	1.57	0.52
39	515,484	46,126	515,915	427	0.08	0.00	0.96	0.26
40	TL	108,000	521,059	637	–	–	–	–
41	475,824	41,669	480,384	1423	0.96	0.96	3.10	2.03
42	TL	108,000	503,238	1,624	–	–	–	–
43	535,369	36	535,369	114	0.00	0.00	1.05	0.21
44	564,517	1173	564,839	195	0.06	0.00	0.94	0.22
45	515,660	93,477	539,932	506	4.71	0.00	4.71	1.23
46	TL	108,000	521,035	765	–	–	–	–
47	TL	108,000	514,500	1729	–	–	–	–
48	TL	108,000	528,860	2081	–	–	–	–

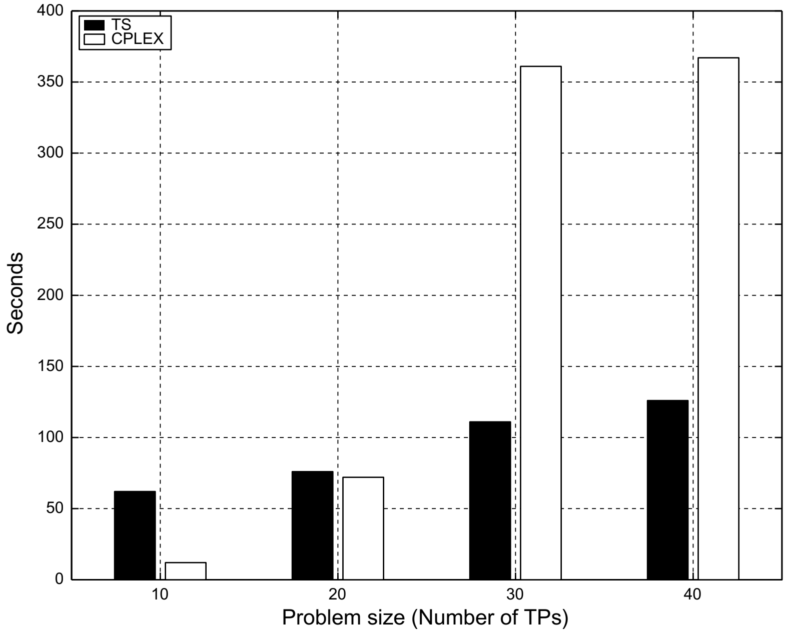

Fig. 3. CPU execution time as a function of the number of TPs.

lowing two columns present the best solution and the CPU execution time found by the TS algorithm. Finally, the last column presents the gap (expressed as a percentage) between the solution value obtained with the TS algorithm and the optimal solution (if the time limit (TL) was not reached).

As can be gathered from Tables 9–11, the proposed TS algorithm is able to find solutions that are, most of the time, quasi-optimal. In fact, this algorithm found the exact solution for 125 instances out of the 191 instances (i.e., 65.45%) for which the gap was evaluated, since the time limit was reached by CPLEX for 49 instances. Moreover, the solutions found are, on average, at 0.56% of the optimal solution, and in the worst case at 6.05%. These results are significantly better than the ones obtain recently by the LS algorithm proposed by St-Hilaire et al. [15] for which the average gap found was 6.53%.

By looking at Tables 9–11, we can see that the CPU execution time of CPLEX varies a lot. This behavior is due to the nature of the branch-and-bound algorithm. Depending on the decision made by the algorithm, shorter or longer CPU time can be necessary to find the solution (i.e., search the tree). However, we can still see a gradual (almost exponential) increase in execution time as the problem size increases. As far as the execution time of the tabu search is concerned, it looks more stable even if we can see a small increase with respect to the problem size.

To compare both methods, we took four problem sizes and, for each one, we randomly generated five instances and calculated the average CPU time for the TS algorithm and for CPLEX. As we can see in Fig. 3, CPLEX is usually faster for small-size instances of the problem. This is a normal behavior since CPLEX is a very powerful tool to solve small-size linear problems. In fact, a few seconds may sometimes be enough to find a solution. Moreover, even if the problem is small, the tabu search still has to go through several iterations (even if the best solution was found in the first iterations). However, for larger size problems, tabu search seems much more appropriate. In fact, the CPU time taken by CPLEX seems to increase exponentially with respect to the problem size. Therefore, for large size problems, the computer might even run out of memory before a solution can be found. Tabu search, on the other side, seems to have a linear increase with respect to problem size.

The main advantage of the TS algorithm over CPLEX is the CPU execution time. We have to remember that the tabu search algorithm is a tradeoff between the quality of the solution and execution time.

5. Conclusions

In this paper, we proposed a TS algorithm in order to find "good" solution to the global UMTS network planning problem in the uplink direction. This problem consists in selecting the number, the location and the type of network nodes (including the node Bs, the RNCs, the MSCs and the SGSNs) as well as the interconnections between them. In fact, this problem includes the cell, the access network and the core network planning subproblems.

It was observed that the tabu-based heuristic produced solutions that were, on average, at 0.56% of the optimal solution. These are good results, considering the difficulty of the problem that we have studied.

There are several avenues of research that are open at this point. First, the only impediment to assess the performance of our method to very large problems lies in the difficulty of evaluating the optimal solution. Such an evaluation becomes time-consuming as the problem instances increase in size. Because the tabu search approach presents relatively stable behavior in terms of closeness to the optimal solution (as demonstrated by the results of the previous section), it could be assumed that such behavior would be maintained for larger instances. The only way to verify this statement is to develop exact methods or efficient ways of evaluating a lower bound. Another research avenue is to develop models and algorithms for the expansion problem of UMTS networks.

Appendix A. Model formulation

A.1. Notation

The following notation used throughout the paper is composed of sets, decision variables, traffic variables, cost parameters and constants.

A.1.1. Sets

- H, the set of TPs;
  - μ_h^c, the number of connections of class $c \in C$ required from TP h;
- C, the set of classes of service (for instance, let $C = \{c_0, \ldots, c_{|C|-1}\}$ such that c_0 is the class for voice service and c_1 to $c_{|C|-1}$ the class for the data services);
  - κ^c is the rate (in b/s) of a connection of class c;
- M_{12},M_{23},M_{24}, respectively the set of links and interface types that can be used to connect the impediment node Bs to the RNCs, the RNCs to the MSCs and the RNCs to the SGSNs;
  - ω^m, the capacity (in b/s) of the link/interface of type m;
  - η^m, the capacity (in circuit) of the link/interface of type m;
- S_1,S_2,S_3,S_4, respectively the set of potential sites to install the node Bs, the RNCs, the MSCs and the SGSNs;
- T_1,T_2,T_3,T_4, respectively the set of node B types, RNC types, MSC types and SGSN types;
  - α^t, the switch fabric capacity (in circuit) of a node of type t;
  - β^t, the switch fabric capacity (in b/s) of a node of type t;
  - n_0^t, the maximum number of interfaces that can be installed in a node of type t;
  - n_1^t, n_2^t, n_3^t, respectively the maximum number of interfaces that can be installed in a node of type t to connect the node Bs, the RNCs and the MSCs/SGSNs.

A.1.2. Decision variables

- v_{01}^{hi}, a 0–1 variable such that $v_{01}^{hi} = 1$ if and only if TP $h \in H$ is connected to a node B installed at site $i \in S_1$;
- v_{12}^{ij}, a 0–1 variable such that $v_{12}^{ij} = 1$ if and only if the node B installed at site $i \in S_1$ is connected to the RNC installed at site $j \in S_2$;
- v_{12}^{ijm}, the number of links of type $m \in M_{12}$ connecting the node B installed at site $i \in S_1$ to the RNC installed at site $j \in S_2$;
- v_{23}^{jk}, a 0–1 variable such that $v_{23}^{jk} = 1$ if and only if the RNC installed at site $j \in S_2$ is connected to the MSC installed at site $k \in S_3$;
- v_{23}^{jkm}, the number of links of type $m \in M_{23}$ connecting the RNC installed at site $j \in S_2$ to the MSC installed at site $k \in S_3$;
- v_{24}^{jl}, a 0–1 variable such that $v_{24}^{jl} = 1$ if and only if the RNC installed at site $j \in S_2$ is connected to the SGSN installed at site $l \in S_4$;
- v_{24}^{jlm}, the number of links of type $m \in M_{24}$ connecting the RNC installed at site $j \in S_2$ to the SGSN installed at site $l \in S_4$;
- x_1^{it}, a 0–1 variable such that $x_1^{it} = 1$ if and only if a node B of type $t \in T_1$ is installed at site $i \in S_1$;
- x_2^{jt}, a 0–1 variable such that $x_2^{jt} = 1$ if and only if an RNC of type $t \in T_2$ is installed at RNC site $j \in S_2$;
- x_3^{kt}, a 0–1 variable such that $x_3^{kt} = 1$ if and only if an MSC of type $t \in T_3$ is installed at MSC site $k \in S_3$;
- x_4^{lt} be a 0–1 variable such that $x_4^{lt} = 1$ if and only if an SGSN of type $t \in T_4$ is installed at site $l \in S_4$.

A.1.3. Traffic variables

- f_{12}^{cij}, the traffic (in circuit) of class $c \in C$ on the link from the node B installed at site $i \in S_1$ to an RNC installed at site $j \in S_2$;
- f_{23}^{cjk}, the traffic (in circuit) of class $c \in C$ on the link from the RNC installed at site $j \in S_2$ to an MSC installed at site $k \in S_3$;
- f_{24}^{cjl}, be the traffic (in circuit) of class $c \in C$ on the link from RNC installed at site $j \in S_2$ to an SGSN installed at site $l \in S_4$.

A.1.4. Cost parameters

- a_{12}^{ijm}, the link and interface costs (including installation cost) for connecting node B installed at site $i \in S_1$ to an RNC installed at site $j \in S_2$ through a link and interface of type $m \in M_{12}$;
- a_{23}^{jkm}, the link and interface costs (including installation cost) for connecting an RNC installed at site $j \in S_2$ to an MSC installed at site $k \in S_3$ through a link and interface of type $m \in M_{23}$;
- a_{24}^{jlm}, the link and interface costs (including installation cost) for connecting an RNC installed at site $j \in S_2$ to an SGSN installed at site $l \in S_4$ through a link and interface of type $m \in M_{24}$;
- $b_1^t, b_2^t, b_3^t, b_4^t$, respectively the cost (including installation cost) of a node B of type $t \in T_1$, of an RNC of type $t \in T_2$, of an MSC of type $t \in T_3$ and of an SGSN of type $t \in T_4$.

A.1.5. Constants

- $P_{\max}$, the maximum emission power of a mobile terminal in TP $h \in H$;
- $P_{\mathrm{target}(t)}$, the minimum power that must be received at the node B of type $t \in T_1$;
- $\mathrm{SIR}_{\min}$, the minimum signal quality that must be received at the node B;
- G, the propagation matrix where $G = [g_{hi}]$ for all $h \in H$ and $i \in S_1$.

A.2. Cost function

The cost function, representing the total cost of the network, is composed of two terms: the cost of the links and interfaces and the cost of the node Bs, RNCs, MSCs and SGSNs.

The cost of the links and interfaces, noted C_{L}, is given by the following equation:

$$C_{\mathrm{L}}(\mathbf{v}) = \sum_{i \in S_1} \sum_{j \in S_2} \sum_{m \in M_{12}} a_{12}^{ijm} v_{12}^{ijm} + \sum_{j \in S_2} \sum_{k \in S_3} \sum_{m \in M_{23}} a_{23}^{jkm} v_{23}^{jkm} + \sum_{j \in S_2} \sum_{l \in S_4} \sum_{m \in M_{24}} a_{24}^{jlm} v_{24}^{jlm} \tag{A.1}$$

The cost of the nodes, noted C_{N}, is given by the following equation:

$$C_{\mathrm{N}}(\mathbf{x}) = \sum_{t \in T_1} b_1^t \sum_{i \in S_1} x_1^{it} + \sum_{t \in T_2} b_2^t \sum_{j \in S_2} x_2^{jt} + \sum_{t \in T_3} b_3^t \sum_{k \in S_3} x_3^{kt} + \sum_{t \in T_4} b_4^t \sum_{l \in S_4} x_4^{lt} \tag{A.2}$$

A.3. The model

The model for the global planning problem of UMTS networks, denoted GPU, can now be given.

GPU:

$$\min(C_{\mathrm{L}}(\mathbf{v}) + C_{\mathrm{N}}(\mathbf{x})) \tag{A.3}$$

subject to:

Node B-type uniqueness constraints

$$\sum_{t \in T_1} x_1^{it} \leqslant 1 \quad (i \in S_1) \tag{A.4}$$

RNC-type uniqueness constraints

$$\sum_{t \in T_2} x_2^{jt} \leqslant 1 \quad (j \in S_2) \tag{A.5}$$

MSC-type uniqueness constraints

$$\sum_{t \in T_3} x_3^{kt} \leqslant 1 \quad (k \in S_3) \tag{A.6}$$

SGSN-type uniqueness constraints

$$\sum_{t \in T_4} x_4^{lt} \leqslant 1 \quad (l \in S_4) \tag{A.7}$$

TP assignment constraints

$$\boxed{\text{Uplink or downlink TP assignment constraints or both}} \tag{A.8}$$

Node B assignment constraints

$$\sum_{j\in S_2} v_{12}^{ij} = \sum_{t\in T_1} x_1^{it} \quad (i \in S_1) \tag{A.9}$$

RNC assignment constraints

$$\sum_{k\in S_3} v_{23}^{jk} = \sum_{t\in T_2} x_2^{jt} \quad (j \in S_2) \tag{A.10}$$

$$\sum_{l\in S_4} v_{24}^{jl} = \sum_{t\in T_2} x_2^{jt} \quad (j \in S_2) \tag{A.11}$$

Node B capacity constraints (at the interface level)

$$\sum_{m\in M_{12}} \sum_{j\in S_2} v_{12}^{ijm} \leqslant \sum_{t\in T_1} n_0^t x_1^{it} \quad (i \in S_1) \tag{A.12}$$

Node B capacity constraints (at the switch fabric level)

$$\sum_{c\in C} \sum_{h\in H} \mu_h^c v_{01}^{hi} \leqslant \sum_{t\in T_1} \alpha^t x_1^{it} \quad (i \in S_1) \tag{A.13}$$

$$\sum_{c\in C} \kappa^c \sum_{h\in H} \mu_h^c v_{01}^{hi} \leqslant \sum_{t\in T_1} \beta^t x_1^{it} \quad (i \in S_1) \tag{A.14}$$

RNC capacity constraints (at the interface level)

$$\sum_{m\in M_{12}} \sum_{i\in S_1} v_{12}^{ijm} \leqslant \sum_{t\in T_2} n_1^t x_2^{jt} \quad (j \in S_2) \tag{A.15}$$

$$\sum_{m\in M_{23}} \sum_{k\in S_3} v_{23}^{jkm} + \sum_{m\in M_{24}} \sum_{l\in S_4} v_{24}^{jlm} \leqslant \sum_{t\in T_2} n_3^t x_2^{jt} \quad (j \in S_2) \tag{A.16}$$

RNC capacity constraints (at the switch fabric level)

$$\sum_{m\in M_{12}} \eta^m \sum_{i\in S_1} v_{12}^{ijm} \leqslant \sum_{t\in T_2} \alpha^t x_2^{jt} \quad (j \in S_2) \tag{A.17}$$

$$\sum_{m\in M_{12}} \omega^m \sum_{i\in S_1} v_{12}^{ijm} \leqslant \sum_{t\in T_2} \beta^t x_2^{jt} \quad (j \in S_2) \tag{A.18}$$

MSC capacity constraints (at the interface level)

$$\sum_{m\in M_{23}} \sum_{j\in S_2} v_{23}^{jkm} \leqslant \sum_{t\in T_3} n_2^t x_3^{kt} \quad (k \in S_3) \tag{A.19}$$

MSC capacity constraints (at the switch fabric level)

$$\sum_{m\in M_{23}} \eta^m \sum_{j\in S_2} v_{23}^{jkm} \leqslant \sum_{t\in T_3} \alpha^t x_3^{kt} \quad (k \in S_3) \tag{A.20}$$

SGSN capacity constraints (at the interface level)

$$\sum_{m\in M_{24}} \sum_{j\in S_2} v_{24}^{jlm} \leqslant \sum_{t\in T_4} n_2^t x_4^{lt} \quad (l \in S_4) \tag{A.21}$$

SGSN capacity constraints (at the switch fabric level)

$$\sum_{m\in M_{24}} \theta^m \sum_{j\in S_2} v_{24}^{jlm} \leqslant \sum_{t\in T_4} \beta^t x_4^{lt} \quad (l \in S_4) \tag{A.22}$$

Node B–RNC link capacity constraints

$$\sum_{c\in C} \kappa^c f_{12}^{cij} \leqslant \sum_{m\in M_{12}} \omega^m v_{12}^{ijm} \quad (i \in S_1, j \in S_2) \tag{A.23}$$

RNC–MSC link capacity constraints

$$\sum_{c\in C} f_{23}^{cjk} \leqslant \sum_{m\in M_{23}} \eta^m v_{23}^{jkm} \quad (j \in S_2, k \in S_3) \tag{A.24}$$

RNC–SGSN link capacity constraints

$$\sum_{c\in C}\kappa^c f_{24}^{cjl} \leqslant \sum_{m\in M_{24}}\theta^m v_{24}^{jlm} \quad (j\in S_2, l\in S_4) \tag{A.25}$$

Traffic flow conservation constraints

$$\boxed{\text{Traffic flow conservation constraints}} \tag{A.26}$$

Additional constraints

$$v_{12}^{ij} \leqslant \sum_{m\in M_{12}} v_{12}^{ijm} \quad (i\in S_1, j\in S_2) \tag{A.27}$$

$$v_{12}^{ij}\max_{t\in T_1}\{n_0^t\} \geqslant \sum_{m\in M_{12}} v_{12}^{ijm} \quad (i\in S_1, j\in S_2) \tag{A.28}$$

$$v_{23}^{jk} \leqslant \sum_{m\in M_{23}} v_{23}^{jkm} \quad (j\in S_2, k\in S_3) \tag{A.29}$$

$$v_{23}^{jk}\max_{t\in T_2}\{n_3^t\} \geqslant \sum_{m\in M_{23}} v_{23}^{jkm} \quad (j\in S_2, k\in S_3) \tag{A.30}$$

$$v_{24}^{jl} \leqslant \sum_{m\in M_{24}} v_{24}^{jlm} \quad (j\in S_2, l\in S_4) \tag{A.31}$$

$$v_{24}^{jl}\max_{t\in T_2}\{n_3^t\} \geqslant \sum_{m\in M_{24}} v_{24}^{jlm} \quad (j\in S_2, l\in S_4) \tag{A.32}$$

Nonnegativity constraints

$$\mathbf{f}\in \mathbf{R}_+^{|C|(|S_1||S_2|+|S_2|(|S_3|+|S_4|))} \tag{A.33}$$

Integrality constraints

$$\mathbf{v}\in \mathbf{B}^{|H||S_1|+|S_1||S_2|+|S_2||S_3|+|S_2||S_4|} \tag{A.34}$$

$$\mathbf{v}\in \mathbf{N}^{|S_1||S_2|(1+|M_{12}|)+|S_2||S_3|(1+|M_{23}|)+|S_2||S_4|(1+|M_{24}|)} \tag{A.35}$$

$$\mathbf{x}\in \mathbf{B}^{|S_1||T_1|+|S_2||T_2|+|S_3||T_3|+|S_4||T_4|} \tag{A.36}$$

The objective function (A.3), as mentioned before, is composed of two terms representing, respectively the cost of the links and the cost of the nodes. Node B-type uniqueness constraints (A.4) impose that at most one type of node B be installed at site $i \in S_1$. Similar uniqueness constraints are necessary for the RNCs, MSCs and SGSNs (see constraints (A.5)–(A.7)). Constraints (A.8) are the TP assignment constraints. Uplink and/or downlink directions can be treated with this model by adding the appropriate constraints. In this paper, we consider the uplink direction. As a result, constraints (A.8) can be replaced by the following set of constraints:

$$\sum_{i\in S_1} v_{01}^{hi} = 1 \quad (h\in H) \tag{A.37}$$

$$v_{01}^{hi} \leqslant \sum_{t\in T_1} x_1^{it} \quad (h\in H, i\in S_1) \tag{A.38}$$

$$v_{01}^{hi} \leqslant \sum_{t\in T_1} \frac{g_{hi}P_{\max}}{P_{\text{target}(t)}} x_1^{it} \quad (h\in H, i\in S_1) \tag{A.39}$$

$$\sum_{t\in T_1} x_1^{it}\left(\sum_{h\in H}\sum_{w\in S_1}\sum_{c\in C}\mu_h^c \frac{g_{hi}}{g_{hw}} v_{01}^{hw} - 1\right) \leqslant \frac{1}{\mathrm{SIR}_{\min}} \quad (i\in S_1) \tag{A.40}$$

Constraints (A.37) impose that each TP $h \in H$ will be assigned to exactly one node B. In other words, each mobile users has to be assigned to one base station. This assignment is also called single homing (as opposed to multi homing where mobile users can be assigned to more than one base station). Constraints (A.38) require that TP $h \in H$ can only be assigned to site $i \in S_1$ if a node B is installed at that site. This makes sure that mobile users will be linked to a location where a base station is installed. Constraints (A.39) impose that TP $h \in H$ can be assigned to site $i \in S_1$ only if the power received at the node B installed at that site is greater or equal to the target value (also called the sensitivity of the base station). Since we are considering the uplink direction, the emitted power of the mobile user equipment minus the propagation loss (between the user and the base station) must be greater or equal than the sensitivity of the base station. Finally, constraints (A.40) are the signal quality constraints. This constraints is summing up intercell and intracell interference so that the signal received at the base station will reach the minimum signal quality required. These constraints are equivalent to the following linear constraints:

$$\mathrm{SIR}_{\min}\left(\sum_{h\in H}\sum_{w\in S_1}\sum_{c\in C}\mu_h^c\frac{g_{hi}}{g_{hw}}v_{01}^{hw}-1\right)\leqslant 1+N\left(1-\sum_{t\in T_1}x_1^{it}\right) \tag{A.41}$$

where N is a large constant.

Constraints (A.9) are node B assignment constraints that impose each node B to be connected to exactly one RNC and constraints (A.10) and (A.11) are the RNC assignment constraints that require each RNC to be connected to exactly one MSC and exactly one SGSN. Constraints (A.12)–(A.22) are the node capacity constraints. It is important to note that the capacity constraints (at the switch fabric level) have been duplicated (in b/s and in circuits) in order to represent the reality of network equipment. For example, since the MSCs handle phone calls (which are best represented in circuits) the capacity of the MSC is given in terms of circuits. Similarly, since the SGSNs are dealing with the data traffic, the capacity is best suited in terms of b/s. Constraints (A.23)–(A.25) are the link capacity constraints. Constraints (A.26), detailed below, are the traffic flow conservation constraints:

$$\sum_{h\in H}\mu_h^c v_{01}^{hi}=\sum_{j\in S_2}f_{12}^{cij}\quad (c\in C, i\in S_1) \tag{A.42}$$

$$\sum_{i\in S_1}f_{12}^{cij}=\sum_{k\in S_3}f_{23}^{cjk}\quad (c=c_0, j\in S_2) \tag{A.43}$$

$$\sum_{i\in S_1}f_{12}^{cij}=\sum_{l\in S_4}f_{24}^{cjl}\quad (c\in C\setminus\{c_0\}, j\in S_2) \tag{A.44}$$

Since the link variables are related to each other, additional constraints (A.27)–(A.32) are required. These constraints are used to bridge the link variables. For example, if there exists a link between a node B and an RNC (i.e., $v_{12}^{ij}=1$), then the total number of links between these two locations must be at least greater or equal to one (i.e., $v_{12}^{ijm}\geqslant 1$). Similarly, if two links are used to link a node B to an RNC (i.e., $v_{12}^{ijm}=2$), then it means that a link must exist between these two locations (i.e., $v_{12}^{ij}=1$). More specifically, constraints (A.27) and (A.28) impose that $v_{12}^{ij}=1$ for all $i\in S_1$ and $j\in S_2$ if and only if $\sum_{m\in M_{12}}v_{12}^{ijm}>1$. Similar constraints are necessary for the links between the sites in S_2 and the sites in S_3 (constraints (A.29) and (A.30)) and for the links between the sites in S_2 and the sites in S_4 (constraints (A.31) and (A.32)). Constraints (A.33) are the nonnegativity constraints which ensure that the flow on the links will always be a positive real value. Finally, constraints (A.34)–(A.37) are the integrality constraints. They guarantee that the variables ($v_{01}^{hi}, v_{12}^{ij}, v_{23}^{jk}, v_{24}^{jl}, x_1^{it}, x_2^{jt}, x_3^{kt}$ and x_4^{lt}) can only take the binary value 0 or 1 and that the variables $v_{12}^{ijm}, v_{23}^{jkm}, v_{24}^{jlm}$ can only take integer values.

GPU is composed of the following three subproblems: the cell planning problem, the access network planning problem and the core network planning problem. Each of these subproblems have already been demonstrated to be NP-hard (see [1,9,10]). As a result, GPU is NP-hard.

References

[1] Amaldi E, Capone A, Malucelli F. Planning UMTS base station location: optimization models with power control and algorithms. IEEE Trans Wireless Commun 2003;2(5):939–52.

[2] Amaldi E, Capone A, Malucelli F. Optimization models and algorithms for downlink UMTS radio planning. In: IEEE wireless communications and networking conference; 2003. p. 827–31.

[3] Blum C, Roli A. Metaheuristics in combinatorial optimization: overview and conceptual comparison. ACM Comput Surv 2003;35(3):268–308.

[4] Chamberland S. An efficient heuristic for the expansion problem of cellular wireless networks. Comput Oper Res 2004;31(11):1769–91.

[5] Chamberland S, Pierre S. On the design problem of cellular wireless networks. Wireless Networks 2005;11(4):489–96.

[6] COST 231 Final Report. Digital mobile radio towards future generation systems; 1999. <http://www.lx.it.pt/cost231>.

[7] Glover F. Tabu search 1. ORSA J Comput 1989;2(1):4–32.

[8] Glover F. Tabu search II. ORSA J Comput 1990;1(3):190–206.

[9] Harmatos J. Planning of UMTS core networks. In: 13th IEEE international symposium on personal, indoor and mobile radio communications; 2002. p. 740–4.

[10] Harmatos J, Juttner A, Szentesi A. Cost-based UMTS transport network topology optimization. In: International conference on computer communication; 1999.

[11] Holma H, Toskala A. WCDMA for UMTS: radio access for third generation mobile communications. John Wiley & Sons Inc.; 2000.

[12] ILOG, Inc.. Using the CPLEX Callable Library and CPLEX Mixed Integer Library. Incline Village: ILOG, Inc.; 2000.

[13] Jonker R, Volgenant T. A shortest augmenting path algorithm for dense and sparse linear assignment problems. Computing 1987;38(4):325–40.

[14] Lauther U, Winter T, Ziegelmann M. Proximity graph based clustering algorithms for optimized planning of UMTS access network topologies. In: IEEE international conference on telecommunications; 2003. p. 1329–34.

[15] St-Hilaire M, Chamberland S, Pierre S. Uplink UMTS network design – an integrated approach. Comput Networks 2006;50(15):2747–61.

[16] Thiel SU, Giuliani P, Ibbetson LJ, Lister D. An automated UMTS site selection tool. In: Third international conference on 3G mobile communication technologies; 2002. p. 69–73.

[17] Wu Y, Pierre S. Optimization of access network design in 3G networks. In: Canadian conference on electrical and computer engineering; 2003. p. 781–4.

[18] Wu Y, Pierre S. A new hybrid constraint-based approach for 3G network planning. IEEE Commun Lett 2004;8(5):277–9.

Marc St-Hilaire received the Computer Engineering degree from the Concordia University in 2002 and the master and Ph.D. degrees in Computer Engineering, respectively in 2003 and 2006 from the École Polytechnique de Montréal. Dr. St-Hilaire is currently an Assistant Professor at the School of Information Technology of the Carleton University. His research interests include wireless network planning, mathematical modeling and metaheuristics.

Steven Chamberland received the Electrical Engineering degree from the École Polytechnique de Montréal, Canada, in 1994 and the Ph.D. degree in Operations Research in 1998 also from the École Polytechnique. From 1998 to 1999, he was with the network architecture group of Bell Canada and from 1999 to 2001, he was a professor at the École de technologie supérieure. Since 2001, he has been an Assistant Professor, an Associate Professor and since 2007, a Full Professor of Computer Engineering at the École Polytechnique de Montréal where he is Director of the Bell Canada Research Chair in Network Architecture. His research interests include telecommunication network planning and network architecture. Dr. Chamberland is a member of IEEE and ACM. He also serves on the Editorial Board of *Computers and Operations Research* and as an Associate Editor for *Telecommunication Systems* and *Information Systems and Operational Research* (INFOR).

Samuel Pierre received the B.Eng. degree in Civil Engineering in 1981 from the École Polytechnique de Montréal, the B.Sc. and M.Sc. degrees in mathematics and computer science in 1984 and 1985, respectively, from the Université du Québec, the M.Sc. degree in economics in 1987 from the Université de Montréal and, the Ph.D. degree in Electrical Engineering in 1991 from the École Polytechnique de Montréal. Dr. Pierre is currently a Professor of Computer Engineering at École Polytechnique de Montréal where he is Director of the Mobile Computing and Networking Research Laboratory (LARIM) and NSERC/Ericsson Industrial Research Chair in Next-generation Mobile Networking Systems. He is the author or co-author of 6 books, 13 book chapters, 16 edited books, as well as over 300 other technical publications including journal and proceedings papers. He received the Best Paper Award of the *Ninth International Workshop in Expert Systems and their Applications* (France, 1989), Distinguished Paper Award from OPNETWORK'2003 (Washington, USA), special mention from *Telecoms Magazine* (France, 1994) for one of his co-authored books, *Télécommunications et transmission de données* (Eyrolles, 1992), among others. His research interests include wireline and wireless networks, mobile computing, performance evaluation, artificial intelligence, and electronic learning. He is a Fellow of Engineering Institute of Canada, senior member of IEEE, a member of ACM and IEEE Communications Society. He is a Fellow of Engineering Institute of Canada, senior member of IEEE, a member of ACM and IEEE Communications Society. He is an Associate Editor of *IEEE Communications Letters* and *IEEE Canadian Review*, a Regional Editor of *Journal of Computer Science*, and he serves on the editorial board of *Telematics and Informatics* published by Elsevier Science.