AN ENUMERATION ALGORITHM FOR OPTIMAL CLUSTERING OF BASE STATIONS IN MOBILE RADIO COMMUNICATION SYSTEM

Syed Zahid Ali *

In a mobile radio communication system, mobile terminals are free to travel within the service area. In order to track a mobile terminal in a cellular mobile network, the radio coverage area of a network is partitioned into clusters of Base Stations, called Location Areas. As mobile terminals cross the boundaries of Location Areas, a significant overhead 'location-updating traffic' is injected into the controlling signalling network. The aim of this paper is to solve the problem of finding optimum Location Areas for a given network such that the load of location-update-signalling traffic on the signalling network is minimized on the basis that the intra-cluster communication is less expensive than the inter-cluster communications. This problem belongs to the family of Non-Polynomial Complete (NP) problems. In this paper, this problem is formulated in terms of an Integer Programming problem and a special tailored Branch and Bound algorithm is applied to find guaranteed optimal solutions. By considering a range of network problems, it is demonstrated that the proposed technique can be applied to solve optimally the real-world network problems

K e y w o r d s: networks, graph theory, optimization, communication system signalling

1 INTRODUCTION

In a mobile communications network, whenever there is a need to establish a communication channel with any particular Mobile Terminal (MT), the network first has to find out which one of the Base Stations (BS) can communicate with this MT. In order to track a MT, the coverage area of a mobile network is partitioned into clusters of BS's or cells. In GSM terminology, these clusters are referred to as Location Areas (LA) [1,2,3]. Current schemes for location management of MT's are based on a two level data hierarchy such that two types of databases, the Home Location Register (HLR) and the Visitor Location Register (VLR), are involved in tracking a MT [3]. The associated network management functions are achieved by the exchange of signalling messages through a Common Channel Signalling System based SS7 signalling network. The mechanism to provide a VLR or a HLR with up-to-date location information relating to each MT is called *Location Updating* (LU) [3]. A MT performs a location update through the SS7 signalling network when it enters a new LA. The higher the inter-LA movement of MT's in a given network, the higher will be the corresponding LU load on the associated signalling network. The call delivery to a MT is achieved by a paging procedure. On receiving the paging signal, the MT sends a reply that allows the associated Mobile Switching Center (MSC) to determine its current residing cell. The traffic generated in a Location Area for tracking a MT is referred to as Paging Traffic (PT) [3].

In order to establish successfully a connection between two end-users, the LU signalling messages may have to go through several intermediate Signalling Transfer Points (STP's) before reaching their destinations. As a consequence, the LU may generate a significant traffic load on the air interface and the SS7 network. Wireless bandwidth is inherently a scarce resource, and the number of radio channels available for signalling is very limited. Therefore, the fast growth in mobile user population may increase proportionally the delay for completing a LU. In order to support a large number of simultaneous connections in a mobile radio network with a smaller overhead traffic on the air interface and on the signalling network, it is imperative to develop sophisticated tools to minimize the traffic on the air-interface and the associated SS7 network. The problem of minimizing this signalling traffic with a paging traffic threshold as a constraint on the size of a Location Area is addressed in this paper and is referred to as the Location Area Design Problem (LADP).

Previous solutions to the problem mainly have been heuristic. In [4], several heuristics were proposed for clustering of BS's in highway and hexagonal cellular systems. In [2] a greedy heuristic is proposed for the design of Location Areas. In [5], a matrix-decomposition based hierarchical algorithm is proposed for solving the LADP efficiently using minimal computational resources. In [6], an algorithm based on generating cutting planes was used for designing Location Areas, which in the case of failure, switches to a Branch and Bound (B&B) algorithm. From the above literature survey it is clear that the LADP has

* Intelligent and Interactive Systems Research Group, Imperial College of Science, Technology, and Medicine, South Kensington, Exhibition Road, London SW7 2BT, England, e-mail: s.zahid@ic.ac.uk

ISSN 1335-3632

been studied by many researchers. However, most of the techniques investigated in the survey are based on heuristic approaches. These heuristic techniques do not provide the guarantee for the global optimality of the yielded solution. These heuristic techniques depend heavily on finding the correct parameters for faster convergence to a solution. In addition, these heuristic techniques are inherently vulnerable to getting trapped in the local optima. The aim of the work reported in this paper is to investigate the viability of using exact solution techniques for finding the truly optimal solution for the LADP. Comparison in terms of computational efficiency, such as CPU time, between the already reported heuristics and the analytical approach presented in this paper, is not the main aim of this paper. In contrast to previously reported results, the present work is of fundamental importance since it provides the first known general analytical solution approach to the LADP, based on a mathematical programming formulation and an associated exact solution technique.

The structure of this paper is as follows: in the Section 2, a formal statement of the problem is presented. In Section 3, the LADP is developed as a node capacitated graph-partitioning problem (GPP). In Section 4, the graph-partitioning problem (GPP) developed is expressed in terms of an Integer Programming (IP) problem. The solution strategy is presented in Section 5. Computational results are given in Section 6. Finally, some conclusions are presented in Section 7.

2 THE PROBLEM STATEMENT

In the solution of LADP, the amount of Paging traffic has to be balanced with the LU traffic. In general, the Paging traffic seems to be less critical for the optimal operation of a GSM network than the traffic caused by LU [2]. A high load of LU traffic in a MSC area causes a high consumption of the signalling resources, stresses the MSC itself, and its related location registers, especially the VLR. Consequently, in this paper, Paging traffic is used as a controlling parameter on the size of each LA while LU traffic is minimized. However, for minimization of LU traffic, the following potential objectives [5] can be considered:

- *Mini-Max* related to the cells; *ie* minimization of the expected LU-traffic that a cell maximally has to handle.
- *Mini-Max* related to Location Areas; *ie* minimization of the expected LU traffic, which maximally has to be handled by the cells of a LA.
- *Mini-Sum* related to the MSC areas. i.e. minimization of the sum of expected LU traffic in the whole MSC area.

In the paper, the LADP is developed as an optimization problem, constrained by upper and lower Paging traffic bounds on the size of each LA, while minimizing the LU traffic for the given network. Consequently, a "*Mini-Sum*" type of optimization objective related to the minimization of sum of expected LU traffic in the whole MSC area is considered. The solution to the LADP partitions a given network into clusters (LA's) of cells such that the inter-cluster movements of MT's are minimized. This will subsequently reduce the LU traffic on the associated SS7 signalling network. Hence the LADP can be stated as follows:

Statement of the Location Area Design problem (LADP)

Given:

(a) MT's inter-cellular movement statistics for a network

(b) the Paging traffic generated for every cell

(c) the Paging traffic thresholds for every Location Area

Determine:

Location Areas such that the LU traffic for a given network is minimized

Subject to:

constraint (c) and other topographical constraints being satisfied.

3 TRANSFORMATION OF THE LADP INTO THE GPP

The LADP can be expressed in terms of a graph-partitioning problem. The cells in a cellular network can be denoted as the nodes of a directed graph, $G = (V, E')$, where $V = \{1, 2, 3, \dots, N\}$ is the set of nodes, E' is the set of directed arcs, and N denotes the number of cells or Base Stations in the network. The amount of PT associated with a cell is represented as a non-negative weight Δv, $v \in V$, of the corresponding node. There is a directed arc $e'_{(i,j)}$ from node v_i to v_j, provided it is possible for a MT to move from i^{th} cell to j^{th} cell. With every directed arc $e'_{(i,j)}$ there is associated a measure for MTs' movement frequency from i^{th} cell to j^{th} cell. This movement frequency of MT's between two adjacent nodes is computed using an appropriate statistical traffic model [6]. These movement frequencies for a given arc are then used to determine associated LU traffic, denoted $c'_{(i,j)}$. In the paper, it is assumed that $c'_{(i,j)}$ and Δv $\forall (i,j)$, $v \in V$ are known parameters.

The directed graph so developed is transformed into a non-directed graph by merging the parallel arcs between each pair of adjacent nodes. Hence, hereafter the notation $e_{(i,j)}$ will refer to a single non-directed arc obtained by merging the arcs between two nodes. The weight of the resultant single arc equals the sum of the constituent arcs *ie* $c_{(i,j)} = c'_{(i,j)} + c'_{(j,i)}$. Although in this paper cellular networks with hexagonal geometry are considered, other geometry can also be dealt with. In this way, the LADP is transformed into a problem involving a planar, non-directed, and weighted graph. In general, the graph developed for a given LADP is not Complete.

As an example, in Fig. 1 a 20-cell mobile network and its associated overlaying graph are shown. The paging requirement for each cell is shown at the center of each cell. The *LU* traffic generated by mobile terminals crossing various cell boundaries is denoted by the weights of the associated arcs.

The LADP therefore can be described as the partitioning of nodes (cells) of a graph (mobile radio network) into required number of clusters (*LA*'s) such that the sum of weights (Paging traffic) of constituent nodes of each cluster is bounded from below by a lower threshold and from above by an upper threshold, while maximizing the sum of costs (LU traffic) on the arcs inside clusters. Maximizing the sum of the costs on arcs within clusters is equivalent to minimizing the sum of costs on the arcs that go between clusters, *ie* the arcs in a *multicut* [7]. The set of arcs whose end-nodes are in different clusters is commonly referred to as the cut-set [7]. From the above discussion it follows that the LADP is essentially a graph-partitioning problem.

When $c_{(i,j)} \geq 0\,,\ \forall (i,j) \in V$, such graph-partitioning problems are referred to as the *Min-Cut Clustering* (MCC) problems [7]. In the language of complexity theory, the MCC problems are Non-Polynomial Complete [7]. Consequently, it follows that the LADP is also a NP-Complete problem.

4 AN INTEGER PROGRAMMING FORMULATION

In this section, the LADP is formulated in terms of an Integer Programming (IP) problem. Let the set

$$\Gamma = \{LA_1, LA_2, LA_3, \dots, LA_M\} \tag{1}$$

be the required partition of graph into M clusters. In cases where M is also an unknown parameter, it is set equal to the number of nodes of the graph, *ie*

$$M = |V|\,. \tag{2}$$

Obviously, the set Γ obeys the following rules:

$$\sum_{i=1}^{M} \text{Cardinality}(LA_i) \leq |V| \ \text{ and } \ \bigcap_{i=1}^{M} LA_i = \{\ \}\,. \tag{3}$$

In order to develop the mathematical model for the LADP, initially two binary decision variables are defined as follows:

The Node Allocation Variables:

The binary variable x_i^k defines the allocation of a node i to a cluster j. It is defined according to the following rule:

$$x_i^k = \begin{cases} 1 & \text{if node } i \text{ is included in cluster } k \\ 0 & \text{Otherwise} \qquad \forall i \in V\,, k \in \Gamma\,. \end{cases} \tag{4}$$

The Arc Allocation Variables:

The variable $z_{(i,j)}^k$ refers to the assignment of an arc $e_{(i,j)}$ to the k^{th} cluster. It is defined according to the following rule:

$$z_{(i,j)}^k = \begin{cases} 1 & \text{if arc } e_{(i,j)} \text{ is included in cluster } k \\ 0 & \text{Otherwise} \qquad (i,j) \in V\,,\ k \in \Gamma\,. \end{cases} \tag{5}$$

Obviously, $z_{(i,j)}^k = 0$ define a set of arcs whose end-nodes belong to different clusters, typically called a *multi-cut* or *cut-set*. An arc $e_{(i,j)}$ can be included in the k^{th} cluster if and only if both of its end nodes are also included in the same cluster. This condition can be formulated by the following constraint:

$$z_{(i,j)}^k \leq \{x_i^k \cdot x_j^k\}\,. \tag{6}$$

The non-linearity of (6) will subsequently render this *IP* problem intractable for most standard solution techniques, such as a combination of *Linear Programming* (*LP*) and *Branch and Bound* (B&B) algorithm. However, it is possible to substitute each such non-linear constraint by a set of equivalent linear inequalities. Hence Inequality (6) is replaced by the following set of constraints:

$$z_{(i,j)}^k \leq x_i^k \qquad \forall (i,j) \in E\,,\ \forall i \in V \tag{7}$$

$$z_{(i,j)}^k \leq x_j^k \qquad \forall (i,j) \in E\,,\ \forall i \in V \tag{8}$$

$$z_{(i,j)}^k \geq x_i^k + x_j^k - 1 \qquad \forall (i,j) \in E\,,\ \forall i \in V\,. \tag{9}$$

Each node must be assigned to *one and only one* cluster. This condition is formulated using the following constraint:

$$\sum_{k=1}^{M} x_i^k = 1\,,\ \forall i \in V\,. \tag{10}$$

Each cluster must contain at least one node to qualify for a valid cluster. This condition can be expressed by the following constraint:

$$\sum_{i \in V} x_i^k \geq 1\,, \quad k = 1, 2, 3, \dots, M\,. \tag{11}$$

In order to constrain the size of each cluster by an upper bound (Paging traffic threshold $P_{\max}$) the following inequality can be written:

$$\sum_{i \in V} \Delta_i \cdot x_i^k \leq P_{\max}\,,\ k = 1, 2, 3, \dots, M\,. \tag{12}$$

A lower bound ($P_{\min}$) is applied on the size of each cluster by the following constraint:

$$\sum_{i \in V} \Delta_i \cdot x_i^k \geq P_{\min}\,,\ k = 1, 2, 3, \dots, M\,. \tag{13}$$

In the case of the LADP, the *LU* traffic measures associated with arcs are strictly non-negative. Therefore, (9) can be dropped, and the optimal solution would automatically satisfy (9).

Let Φ be the value of *Minimum-Cut*. Hence objective function can be defined as follows:

$$\Phi_{z^k_{(i,j)}} = \min\Big(\sum_{(i,j)\in E} c_{(i,j)} - \sum_{k=1}^{M} \sum_{(i,j)\in E} c_{(i,j)} \cdot z^k_{(i,j)} \Big) . \quad (14)$$

All the variables (4–5) are binary, the constraints (7–13) and the objective function (14) are linear (in)equalities. Therefore, the above model essentially represents an Integer Programming (*IP*) problem.

5 THE SOLUTION STRATEGY

The *IP* problem developed in the previous section belongs to the family of NP-Complete problems. Until now, no polynomial algorithm is reported that can solve any *NP-Complete* problem with guaranteed optimality. In order to solve this model, a special Branch and Bound (B&B) algorithm is used in this paper. Instead of attempting to solve the problem directly over the set of all feasible solutions, the B&B algorithm successively divides this set into increasingly smaller sets which have the property that any optimal solution must be in one of the sets. The essence of this B&B algorithm is the underlying application of Linear Program (*LP*) [8-9]. The convex set defined by the constraints of an *LP* is called a *polyhedron* and a bounded polyhedron is called a *polytope*. *LP* plays an important role in discarding the uninteresting parts of the enumeration tree.

If the *total* solution space of the *IP* is denoted as Sol^1 and that of Integer solutions as Sol^2, then two problem sets P(1) and P(2) can be defined as follows:

$$\text{P(1)}: \Phi_{z^k_{(i,j)}} = \min\Big(\sum_{(i,j)\in E} c_{(i,j)} - \sum_{k=1}^{M} \sum_{(i,j)\in E} c_{(i,j)} \cdot z^k_{(i,j)} \Big) ,$$

$$0 \le \{z^k_{(i,j)}, x^k_i \in Sol^1\} \le 1$$

is said to be a relaxation of

$$\text{P(2)}: \Phi_{z^k_{(i,j)}} = \min\Big(\sum_{(i,j)\in E} c_{(i,j)} - \sum_{k=1}^{M} \sum_{(i,j)\in E} c_{(i,j)} \cdot z^k_{(i,j)} \Big),$$

$$\{z^k_{(i,j)}, x^k_i \in Sol^2\} \in \{1,0\} \text{ if and only if } Sol^1 \supseteq Sol^2 .$$

According to the theory of polyhedron [8–9], if $\Phi^{Opt\_LP}_{z^k_{(i,j)}}$ is an optimal solution to P(1) and $\Phi^{Opt\_IP}_{z^k_{(i,j)}}$ is an optimal solution to P(2), then for the proposed minimization model, it follows

$$\Phi^{Opt\_LP}_{z^k_{(i,j)}} \le \Phi^{Opt\_IP}_{z^k_{(i,j)}} .$$

Furthermore, if $\Phi^{Opt\_LP}_{z^k_{(i,j)}} \in Sol^2$ implies that $\Phi^{Opt\_LP}_{z^k_{(i,j)}}$ is an optimal solution to P(2) as well.

The Simplex [8–9] based B&B algorithm begins by solving the original problem as an *LP* problem by relaxing integrality conditions *ie* $0 \le z^k_{(i,j)}, x^k_i \le 1$. If an integer solution is not obtained, the next step is to pick one of the variables to create two descendants of the original problem, one with $z^k_{(i,j)} = 0$, the other with $z^k_{(i,j)} = 1$. Thus the original problem is replaced by these two descendant sub-problems. The process then repeats and creates a search solution tree. Eventually an integer solution is obtained for one of the current *IP* problems, whereupon this solution provides a candidate for the optimal solution to the original problem. This process continues until each unexamined sub-problem is decomposed, solved, or shown not to be leading to an optimal solution. Keeping track of the current best solution of these descendant problems and its associated objective function value provides an additional way to eliminate descendants of the original *IP* problem. The current best candidate problem is called the incumbent. The possibilities for this type of approach can be itemized as follows:

Let S^* denote the objective function value of the incumbent solution. Then, whenever a current *IP* problem is solved as an *LP* problem, one of the following four alternatives may arise:

(a) The relaxed *LP* problem has no feasible solution (that is, the associated *IP* problem is also infeasible).

(b) The *LP* problem has an optimal solution with $S_0 \le S^*$, *ie* the current *IP* problem optimum must also yield $S_0 \le S^*$.

(c) The optimal solution to the *LP* problem is both integer feasible and yields $S_0 > S^*$ (in which case the solution provides an improved incumbent for original *IP* problem, S^* is set to S_0).

(d) None of the foregoing occurs; *ie* the optimal solution exists, satisfies $S_0 > S^*$, and is not integer feasible.

In cases (a) to (c), the current *IP* problem is disposed off simply by solving the associated *LP* problem. Otherwise the problem requires further exploration. These alternatives help to reduce the search space significantly since the algorithm avoids the consideration of solutions (or branches in the solution tree) that cannot be optimal.

Clique cuts were generated during the enumeration process. In addition, the Cover cuts [9] were generated to strengthen the inequality (12). A rounding heuristic [10] was used at the root node to obtain an initial solution to help faster convergence of the B&B algorithm to optimality. Branching priorities were assigned to different binary variables heuristically. Variables with higher priorities are to be branched on to the variables with lower priorities. For the LADP, these branching priorities were assigned on the analysis of Paging traffic requirements of cells, that is, the nodes with higher weights were assigned higher priority than those with lower weights.

Table 1. Traffic-related specifications for the ladp for the area of seoul

Cell No.	Incoming call arrival rate (calls/h/MT)	Cell size (km^2)	Population density of switched on MT (MT's/km^2)	Paging traffic measure Δ_i (in bits)
1	2	4	30	111120
2	2	4	30	111120
3	2	4	30	111120
4	2	4	30	111120
5	2	4	30	111120
6	2	4	30	111120
7	2	4	30	111120
8	2	4	30	111120
9	3	2.25	66	206266.5
10	3	2.25	66	206266.5
11	3	2.25	66	206266.5
12	3	2.25	66	206266.5
13	2	4	30	111120
14	2	4	30	111120
15	2	4	30	111120
16	3	2.25	66	206266.5
17	2	4	30	111120
18	2	4	30	111120
19	2	4	30	111120
20	2	4	30	111120
21	3	2.25	66	206266.5
22	2.5	2.56	60	177792
23	2.5	2.56	60	177792
24	2.5	2.56	60	177792
25	2.5	2.56	60	177792
26	2.5	2.56	60	177792
27	2.5	2.56	60	177792
28	2.5	2.56	60	177792
29	2.5	2.56	60	177792
30	2.5	2.56	60	177792
31	2.5	2.56	60	177792
32	2.5	2.56	60	177792
33	3	2.25	66	206266.5
34	3	2.25	66	206266.5
35	3	2.25	66	206266.5
36	3	2.25	66	206266.5
37	3	2.25	66	206266.5
38	3	2.25	66	206266.5

Table 2. Specifications of location areas developed for the area of seoul

Location Area No	No. of cells allocated	Cells allocated
1	9	1,2,3,4,5,8,9,10,11
2	7	6,7,27,28,29,30,31
3	12	12,13,14,15,16,17, 18,19,20,21,37,38
4	10	22,23,24,25,26,32,33,34,35,36

In order to exploit some structural features of the problem, a prefixing heuristic was developed to fix some variables. The heuristic identifies the set of those nodes that are so apart geographically that two or more nodes of this set cannot be included together in a cluster. Each node of this set is then assigned to a different and empty cluster.

Advanced pre-processing [10] techniques were applied to the IP formulation before invoking the B&B algorithm. A probing technique [10] was used during the enumeration process.

6 PROBLEM SPECIFICATIONS AND COMPUTATIONAL RESULTS

The proposed methodology has been applied to solve a range of network problems. The computations were carried out on a 800 MHz Pentium-III machine. The mathematical modelling was performed in General Algebraic Modelling Systems [11], and the Cplex [10] was used as LP solver.

Given various parameters and the number of LA to be developed, the objective of solving various problems is to minimize the Location Updating traffic measure. It was stated earlier via (2) that if the number of LA's is not fixed, then the cardinality of the set Γ is set equal to $|V|$. In this case, the solution obtained will yield the optimal number of LA's together with the allocation of cells to their respective LA's.

In order to demonstrate how effectively the proposed method can determine the LA's, consider a LADP for the area of Seoul. The problem is taken from [6] involving one quarter of the city of Seoul and including 38 cells. Following the discussion in [6], the bandwidth cost for a single paging is set to 463 bits. Similarly, the bandwidth cost for a single LU is set to 1231 bits. The incoming call rate in a cell is estimated using the following equation [6]:

$$\lambda_i = \lambda \cdot \rho_i \cdot A_i \, , \; i \in V \qquad (15)$$

where, λ: the arrival rate of incoming call per Mobile Terminal (call/hour/MT); ρ_i: the density of switched on MT's in cell 'i' (MT's/km^2); A_i: the area of i^{th} cell (km^2).

The incoming call rate per MT, cell size, the population density and the paging traffic costs for each cell are listed in Table 1.

The MT's cell crossing rate $c_{(i,j)}$ between cell i and cell j is computed using the following formula:

$$c_{(i,j)} = l_{(i,j)} \cdot \frac{a_{(i,j)}}{b_{(i,j)}} \cdot \rho_{(i,j)} \cdot v_{(i,j)} \qquad (16)$$

where $l_{(i,j)}$: the length of border between cell i and j; $a_{(i,j)}$: the percentage of border length $l_{(i,j)}$ corresponding to streets and pavements; $b_{(i,j)}$: the percentage of the area around border covered by streets and pavements; $\rho_{(i,j)}$: the density of the switched-on MT around border in MT's/km^2; $v_{(i,j)}$: the average speed for MT's moving near border in km/h.

In this problem, the values of $a_{(i,j)}$, $b_{(i,j)}$ and $v_{(i,j)}$ were assumed to be uniform throughout the design area.

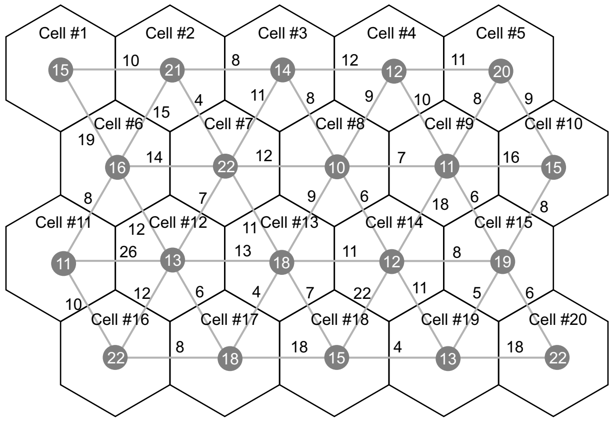

Fig. 1. A 20-nodes network. Numbers inside black circles represent paging traffic measures and values shown on the arcs represent the frequency of intercellular movements of mobile terminals.

Table 3. Problem specifications and computational results

Problem No.	No. of nodes (cells or BS's) in network	No of LA's required	Value of Min-Cut obtained	Computational time (in seconds)
1	20	4	148	324
2	32	5	160	386
3	58	3	528	540
4	95	6	492	566
5	80	5	468	612
6	120	4	562	1012
7	150	9	652	1469
8	300	15	2153	6500

The percentages $a_{(i,j)}$ and $b_{(i,j)}$ were set to 0.6 and 0.3, respectively. Two types of MTs with equal distribution were considered: moving vehicles and pedestrians. Their speeds were assumed to be 15 km/h and 5 km/h respectively. However, the average value of 10 km/h is used in the computations. The MT density was assumed to be uniform in a cell and its value around the border was computed as follows:

$$\rho_{(i,j)} = \frac{\rho_i + \rho_j}{2}\,. \tag{17}$$

The border lengths were assumed to vary from 50 metres to 1 km. Equation (16) quantifies the cell-crossing rate along a border. Since the bandwidth cost of each cell crossing is 1231 bits, the cell crossing rate obtained using (16) is multiplied by 1231 in order to estimate the cost of selecting the corresponding arc in the min-cut. As an example, consider the cell No. 1 and cell No. 2 listed in Table 1. The length of border between cell No. 1 and cell No. 2 is 0.5 km, $a_{(i,j)} = .6$, $b_{(i,j)} = .3$, $v_{(i,j)} = 10$ km/h and $\rho_{(i,j)} = 30\,\mathrm{MT/km}^2$ gives a cell crossing rate of 300 MT's/hour. Therefore, the arc corresponding to this border is assigned a LU traffic cost of $300 \cdot 1231 = 36900$ bits. In this benchmark problem, the number of LA's is also an unknown variable. Therefore, the value of M was set to 38. The value of $P_{\max}$ in (12) was set to 1450000 and the value of $P_{\min}$ in (13) was set to 150000. Using the approach presented, the graph of 38 cells was partitioned into 4 subgraphs (LA's). The value of the associated min-cut was found to be 9793836. The CPU solution time was reported to be 1620 seconds. The long computational time is attributed to the fact that M was set to 38, and this upper bound led to a large number of binary variables. The LA's developed are listed in Table 2.

As another example, a LADP involving 49 cells was selected from [12]. In this example, the hexagonal cellular geometry was considered. The bandwidth cost associated with the cell-crossing of MT's along a border was computed using the following equation:

$$LU_i = \frac{A_i v_i s_i}{\pi} m_u(\rho_o - tr) \tag{18}$$

where A_i: the population density in cell i; v_i: the speed of Mobile Terminals in cell i; s_i: the side length of hexagonal cell; ρ_0: the probability that a Mobile Terminal is switched on; tr: average traffic intensity per MT; m_u: the length of Location update message.

The paging traffic in each cell is computed using the following relationship:

$$\Delta_i = \frac{tr}{\tau_i} A_i 3\sqrt{3} s_i^2 m_p p_m\,. \tag{19}$$

where τ_i: average call holding time; p_m: probability that there is a mobile terminated call; m_p: length of the paging message. In this problem, ρ_0 and p_m were set to 0.5 and 0.3, respectively. m_u and m_p were assumed 18 and 8 bytes, respectively. The average call holding was assumed to be 170 seconds. In this problem, it is given that the network consisting of 49 cells needs to be partitioned into five LA's. Using the proposed approach, the min-cut was found for this problem in 840 seconds with a value of 10953.

In order to further test the proposed methodology, a number of network problems were considered. The set of problems considered and their specifications are presented in Table 3. For each problem, the total number of Base Stations (cells or nodes) in the network, and the number of LA's required are listed. Table 3 also lists the computational results obtained for the set of examined problems. For each problem, the value of the min-cut together with associated solution time is listed.

Problem No. 8 is a big problem involving a network of 300 cells. It is assumed that it is required to partition this network into 15 Location Areas. The paging requirements of BS's were assumed to vary between 3.5 kbits and 14 kbits. The upper bound on the paging traffic of LA was set to 269 kbits; and the lower bound to 25 kbits. The inter-cellular movement frequency (also known as cell-crossing rate) of MT between the adjacent cells was assumed to vary between 0 and 12 MT's/h. Using the proposed solution approach, the problem was solved optimally in 6500 seconds with the min-cut value of 2153.

7 CONCLUSIONS

In order to track the location of a mobile terminal in a cellular mobile network, the radio coverage area of a network is partitioned into clusters of base stations called Location Areas. As MT's cross the boundaries of LA's, a significant overhead location-updating traffic is injected into the controlling signalling network. The intent of this paper was to find guaranteed optimum LA's for a given network such that the load of location-update-signalling traffic on the signalling network is minimized in the context that the intra-cluster communication is less expensive than the inter- cluster communication. The problem is formulated an Integer Programming problem, and a special B&B algorithm is applied to solve it. By considering a range of network problems, it is demonstrated that the proposed exact solution technique can be used to solve fairly large-size network problems with guaranteed optimality of the solutions obtained.

Acknowledgements

The author is thankful to Prof. L. F. Turner of Imperial College of Science, Technology and Medicine, London SW7 2 BT, for providing advise, guidance, computing resources and necessary support for conducting the research reported in this paper.

References

[1] ETSI/TC, "Mobile Application Part (MAP) specifications, version 4.8.0," Tech. Rep. Recommendation GSM 09.02, 1994.

[2] PLEHN, J. : The Design of Location Areas in a GSM Network, 45th IEEE VTC, pp. 871–875, Vol. 2, 25–28 July, 1995.

[3] AKYILDIZ, I. F. : On Location Management for Personal Communications Networks, IEEE Communication Magazine **34**, 138–145, Sept. 1996.

[4] LIM, K.—LEE, Y. H. : Optimal Partitioning of Heterogeneous Traffic Sources in Mobile Communications Network, IEEE Transactions on Computers **46** No. 3, March 1997.

[5] ALI, S. Z. : Design of Location Areas for Cellular Mobile Radio Network, in proc. IEEE Vehicular Technology Conference, vol. 3, pp. 1106–1110, May 6–9, 2002. Birmingham AI, USA.

[6] TCHA, D-W—CHOI, T-J—MYUNG, Y-S : Location Area Partition in a Cellular Radio Network, Journal of the Operational Research Society **48** (1997), 1076–1081.

[7] FERREIRA, C. E.—MARTIN, A.—De SOUZA, C. C.—WEISMANTEL, R.—WOLSEY, L. A. : The Node Capacitated Graph Partitioning Problem: A Computational Study, Mathematical Programming **81** (1998), 229–256.

[8] GARFINKEL, R. S.—NEMHAUSER, G. L. : Integer Programming, Wiley, New York, 1972.

[9] AVRIEL, M.—GOLANY, B. : Mathematical Programming for Industrial Engineers, Marcel Dekker Inc, 270 Madison Avenue, New York 10016, 1996.

[10] CPLEX, IPLOG CPLEX Divison, 889 Alder Avenue, Suite 200, Incline Village, NV 89451 USA..

[11] BROOKE, A.—KENDRICK, D : GAMS: A User's Guide, December 1998, GAMS Development Corporation..

[12] SUNG, C. S.—HONG, J. M. : Optimal Base Station Clustering for a Mobile Communication Network Design, in proc. ITST-2000, Oct 2000.

Received 6 June 2002

Syed Zahid Ali (Dr) received his BE in Electrical Engg. from N.E.D. University of Engg. and Technology, Karachi. He received his MES degree in Electrical Engineering from the Graduate School of Engineering, University of Sydney, Australia. He received his PhD and DIC degrees from the University of London and Imperial College London in Electrical Engineering. In addition, he holds certificate courses in Business Administration from the Institute of Business Administration and Specialist certificate in switching systems from Alcatel Institute-de-training Lanion, France. He has held several technical, research, instructional, operational and managerial posts in Pakistan Telecom, Pakistan and Colt Telecom, Zurich Switzerland, Middlesex University, London and Imperial College, London. His areas of research interest include network design, cellular systems, radio resource allocation, mathematical programming, optimization techniques, graph partitioning techniques, switching systems, operations research and neural networks.