Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research 164 (2005) 301–323

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/dsw

Production, Manufacturing and Logistics

Base station location and channel allocation in a cellular network with emergency coverage requirements

Mohan R. Akella [a,b], Rajan Batta [a,b,c,*], Eric M. Delmelle [a,d], Peter A. Rogerson [a,c,d], Alan Blatt [a], Glenn Wilson [a]

[a] *Center for Transportation Injury Research, General Dynamics, Buffalo, NY 14225, USA*
[b] *Department of Industrial Engineering, University at Buffalo (SUNY), 342, Bell Hall, Buffalo, NY 14260, USA*
[c] *National Center for Geographic Information and Analysis, University at Buffalo (SUNY), Buffalo, NY, USA*
[d] *Department of Geography, University at Buffalo (SUNY), Buffalo, NY, USA*

Received 15 April 2003; accepted 8 December 2003
Available online 28 February 2004

Abstract

The location of base stations (BS) and the allocation of channels are of paramount importance for the performance of cellular radio networks. Also cellular service providers are now being driven by the goal to enhance performance, particularly as it relates to the receipt and transmission of emergency crash notification messages generated by automobile telematics systems. In this paper, a Mixed Integer Programming (MIP) problem is proposed, which integrates into the same model the base station location problem, the frequency channel assignment problem and the emergency notification problem. The purpose of unifying these three problems in the same model is to treat the tradeoffs among them, providing a higher quality solution to the cellular system design. Some properties of the formulation are proposed that give us more insight into the problem structure. An instance generator is developed that randomly creates test problems. A few greedy heuristics are proposed to obtain quick solutions that turn out to be very good in some cases. To further improve the optimality gap, we develop a Lagrangean heuristic technique that builds on the solution obtained by the greedy heuristics. Finally, the performance of these methods is analyzed by extensive numerical tests and a sample case study is presented.

Keywords: Health sciences; Logistics; Location

1. Introduction

Motor vehicle crashes are a major health problem and an economic burden in the United States, see SmartRisk (1998) and Walker (1996). According to the National Highway Traffic Safety Administration

* Corresponding author. Address: Department of Industrial Engineering, University at Buffalo (SUNY), 342, Bell Hall, Buffalo, NY 14260, USA. Tel.: +1-716-645-2357; fax: +1-716-645-3302.
E-mail address: batta@eng.buffalo.edu (R. Batta).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2003.12.014

(NHTSA, 2000), there were over 6.3 million motor vehicle crashes in 1999. These crashes led to over 40,000 deaths. Of note, approximately 50% of the fatalities occur before the crash victim reaches a hospital.

Current research is branching in two directions. One deals with methods of preventing vehicle crashes on roads. The second, and more pertinent to this work, is dealing with ways of reducing the response time in the event of a crash or a fatality. There is a substantial body of literature regarding the impact of emergency medical service (EMS) response time and time to definitive care on trauma victim outcomes. Terms like 'golden hour' in Jacobs et al. (1984) and Lerner and Moscati (2001), 'silver day' in Blow et al. (1999) and 'platinum ten minutes' have been coined to describe the importance of time in treating trauma injuries. Evanco (1999) establishes a quantitative relationship between fatalities and crash notification time. According to this paper, if a rural mayday system were implemented (i.e., a 100% market penetration) and the service availability were 100%, then we would expect monetary benefits of about \$1.83 billion per year and comprehensive benefits (which includes the monetary value attached to the lost quality of life) of \$6.37 billion per year. More recently Clark and Cushing (2002) studied data from fatal crashes to predict the effect of a fully functional ACN system on reducing crash-related mortality in the United States. They estimate that an ideal system would reduce crash fatalities by 2–6% a year.

Location of base stations and channel allocation in cellular communications plays a major role in reducing the notification time in the event of a crash, especially in rural areas where coverage is weak. In this work we address tradeoff issues faced by a cellular service provider who needs to render efficient coverage to both "emergency" as well as "regular" calls.

1.1. Automated Crash Notification (ACN) systems

Emergency Notification and Response (EN & R) systems and associated services aid a specific individual or motorist to request help from, and provide information to, authorities about a distress situation. Crucial to getting adequate help to a crash victim is prompt notification that (a) a crash has occurred, (b) the location of the crash, and (c) some measure of the severity or injury-causing potential of the collision. Automated Crash Notification (ACN) systems capable of performing tasks (a) and (b) have been installed as expensive options on a limited number of high-end luxury cars. These devices are activated by air bag deployment. More advanced sensors can also estimate the injury-producing capability of the crash. The first estimate of the number of potential lives saved by ACN technology is 3000 lives per year according to Champion et al. (1998). In general, there are many reasons that can cause an emergency crash message to fail to be generated or completed, including:

- Damage caused to the ACN device due to the severity of the crash.
- Loss of primary and backup power in the vehicle as a result of the crash.
- Weak signal strength due to poor cellular coverage, damage to vehicle antenna, or final resting position of the crashed vehicle (i.e., rolling into a ditch).
- Insufficient cell channel capacity.

1.2. Base station location, channel allocation and access protocols

The following are the four basic components of a cellular mobile network:

- mobile station,
- base station,

- mobile switching center,
- public switched telephone network.

The mobile station (MS) constitutes the interface between the mobile subscriber and the base station. Base stations are responsible for serving the calls to or from the mobile units located in their respective cells. The mobile switching center (MSC) is a telephone exchange especially assembled for cellular radio services. Finally, the public switched telephone network (PSTN) treats the MSCs as ordinary fixed telephone exchanges.

Given an area to serve the teletraffic, cellular providers would have to decide the following:

- The number of base stations to be located. This would depend on budget limitations.
- Optimal positioning of the base stations to maximize the coverage in the region given a restriction on the number of base stations to be built (particularly true in rural areas where the number of base stations is less and their locations are thus more critical).
- Channel capacity of each base station subject to the total channel capacity. This would mainly depend on the teletraffic demand at each base station.
- Maximal base station transmitting power.
- Antenna height.

The first three aspects constitute the design of the cellular network and are one of the major problems in cellular communications.

A subsequent problem in the design of cellular communications is the efficient use of the limited available radio channels. There are two strategies for assigning channels to cells: Fixed Channel Assignment (FCA), and Dynamic Channel Assignment (DCA). The FCA strategy allocates channels to each cell in advance according to estimated traffic intensity in the cell. The DCA strategy foresees the assignment of radio resources to various cells dynamically in real time, to meet rapidly changing demand for communication channels.

We now review previous work in the optimal positioning of base stations and in channel allocation. The *Adaptive Base Station Positioning Algorithm* (ABPA) was introduced by Fritsch et al. (1995). It uses an early version of the Demand Node Concept and the major drawback of ABPA is its lack of speed. A promising approach to automatic network design was presented by Chamaret et al. (1997). The radio network design task is modeled as a maximum independent set search problem. In contrast to this, Ibbetson and Lopes (1997) proposed an algorithm that considers only traffic distribution as a constraint for cell site locations.

The approach for the design of micro-cellular radio communications proposed by Sherali et al. (1996) concentrates on radio frequency (RF) constraints, since in the considered micro-cellular environment network capacity is not of major importance. They used well established non-linear local optimization algorithms (simplex method, i.e., Hooke and Jeeve's method, quasi-Newton, and conjugate gradient) in evaluating the objective function. Tcha et al. (2000) addressed the radio network design problem in a Code Division Multiple Access (CDMA) system. They use two heuristics: the construction heuristic for choosing an initial feasible subset of potential sites, and the improvement heuristic for reducing the cost associated with the selected subset by changing some of its constituent sites. Wright (1998) employed a direct search method to finding the optimal placement of base stations, since it requires only the value of the function to be optimized. Bose (2001) used dynamic programming to determine the optimal placement of base stations in an urban setting, given the cell coverage. Statamatelos and Ephremides' (1996) objective function was based on maximizing the coverage area while minimizing co-channel interference, and incorporated spatial diversity.

We now present an overview of the cellular protocols that can be used to design a wireless network. Frequency Division Multiple Access (FDMA) assigns individual channels to individual users. These channels are assigned on demand to users who request service. During the period of the call no other user can share the same frequency band. Time Division Multiple Access (TDMA) systems divide the radio spectrum into time slots, and in each slot only one user is allowed to either transmit or receive. In Code Division Multiple Access (CDMA), all users use the same carrier frequency and may transmit simultaneously. Each user has its own pseudorandom codeword. The receiver performs a time correlation operation to detect only the specific desired codeword.

1.3. Coverage models

The assumption underlying all coverage models is that customers beyond a specified service range are not adequately served by the service facilities. The objective of the *Set Covering Problem* (SCP) is to determine the number of required service centers, i.e. base stations, and their locations such that *all* users of the wireless network are served with an adequate service level, i.e. field strength level. However, for an economic design of wireless communication networks, a tradeoff between the cost of coverage and the benefit resulting from covering this area is desired. Church and ReVelle (1974) define this problem as the *Maximum Coverage Location Problem* (MCLP). The MCLP assumes a limited budget and includes this as a constraint on the number of facilities to be located. The book by Daskin (1995) contains a thorough discussion of coverage models and their applications. Our model builds on the SCP and MCLP models in the context of a cellular application.

1.4. Motivation

To reduce crash-related fatalities and minimize crash notification times, NHTSA sponsored Veridian Engineering in the Automated Collision Notification (ACN) Field Operational Test Program from 1995 to 2000. ACN explored the ability of in-vehicle equipment to reliably sense and characterize crashes, and automatically transmit crash location and crash severity data to the proper public safety agencies. The paper by Akella et al. (2003) summarizes these findings. According to the paper, 70 crashes involving ACN-equipped vehicles occurred within Erie County, New York. Of the 22 crashes where the severity level was above the threshold, 14 ACN systems detected the crash and alerted the Erie County Sheriff. The failure to notify the EMS in the remaining 8 crashes can be attributed to insufficient signal strength. This number is quite large and hence to eliminate any possibility of failure of the ACN device to notify due to a weak signal, the Received Signal Strength Indicator (RSSI) should be strong at *potential* crash locations.

Though we seek to cover nodes prone to vehicular crashes, it is not only limited to them. For example, defense establishments, nuclear power stations, high rise buildings etc. which are prone to enemy attacks need to have sufficient coverage. In the event of an emergency, numerous calls would originate from those regions and covering such calls is crucial. Over the past few years the FCC has been taking several important steps in the US (see FCC, 1996) to foster major improvements in the quality and reliability of 911 services available to the customers of wireless telecommunications service providers. We note that this special coverage could also be profit motivated if the provider identifies a group of important customers to whom high quality service is imperative.

This paper makes the following three main contributions:

- It models a typical cellular network design problem from the perspective of emergency notification.

- It introduces a unique formulation involving the MCLP with set covering constraints.
- It proposes efficient heuristic solution techniques for this problem.

2. Model formulation

We use the discrete population model for the traffic description, denoted as the Demand Node Concept (DNC) introduced by Tutschku et al. (1996). A demand node represents the center of an area from which a given number of call requests per unit time originate. To take into account the time variation of call traffic, each day is divided into a fixed number of time slots. We assign a weight to each time slot in order to differentiate the importance attached to the hour of the day when calls are placed. For example, calls placed during the day might be considered more important than those placed during the night, since most business calls are made during the day. It is pointed out that the weights attached to each time slot bear no direct relation to the coverage of emergency nodes. In our model we incorporate mandatory coverage to all the crash nodes. So irrespective of the values of the weights, the crash nodes are always covered. In our computational experiments, we assigned weights randomly due to lack of knowledge of how cell phone companies would assign importance to the time of the day.

An emergency/crash node represents a region that is prone to crashes. Also, there is a limit on the number of available channels per time slot. We are initially given a set of potential locations of base stations. The problem is to find an optimal set of locations of a given number of base stations that would maximally cover the demand nodes based on their demands and cover the emergency nodes. We call this the Network Design Emergency Coverage (NDEC) model. We assume that our network is designed for a homogenous system i.e., it uses a single protocol for transmissions. With this assumption, our model becomes independent of the cellular protocol (CDMA, TDMA, etc.) in use as long as it can be transformed to channel capacity requirement as stated in the formulation. We formulate the problem as a Mixed Integer Programming (MIP) problem. We assume that the demand nodes and the emergency nodes are spatially static with time and that we know the demands of each demand node for all time slots.

2.1. Network Design Emergency Coverage (NDEC) model

Sets

M set of possible locations of base stations,
N set of demand nodes,
E set of emergency nodes.

Constants
T total channel capacity,
p the number of base stations to be located,
W_t importance attached to time slot t,
H_{jt} demand at node j at time t,
A_{ij} $\begin{cases} 1 & \text{if BS } i \text{ covers node } j, \\ 0 & \text{otherwise.} \end{cases}$

Variables
f_{ijt} fraction of demand of node j satisfied by BS i at time slot t,
x_i $\begin{cases} 1 & \text{if there is a base station at location } i, \\ 0 & \text{otherwise.} \end{cases}$

$$(\mathbf{P1}) \quad \text{Maximize} \quad \sum_{i \in M} \sum_{j \in N} \sum_{t} W_t H_{jt} f_{ijt} \tag{1}$$

subject to

$$\sum_{i \in M} x_i \leqslant p, \tag{2}$$

$$f_{ijt} \leqslant A_{ij} x_i \quad \forall i \in M, j \in N, t, \tag{3}$$

$$\sum_{i \in M} f_{ijt} \leqslant 1 \quad \forall j \in N, t, \tag{4}$$

$$\sum_{i \in M} \sum_{j \in N} H_{jt} f_{ijt} \leqslant T \quad \forall t, \tag{5}$$

$$\sum_{i \in M} A_{ik} \cdot x_i \geqslant 1 \quad \forall k \in E, \tag{6}$$

$$x_i \in \{0, 1\} \quad \forall i \in M,$$

$$f_{ijt} \geqslant 0 \quad \forall i \in M, j \in N, t.$$

The A_{ij} matrix defined above is a 0–1 matrix that indicates whether a node at j can be covered by a BS at i. Note that distance might not be the only criterion for coverage. Obstructions from buildings, multiple reflections on walls etc. affect the signal strength at any point. We assume that the A_{ij} matrix has been constructed taking into account these factors. The total channel capacity is assumed to be the same for all times slots. The objective function (1) maximizes the demand coverage over all time slots in a day. Constraint (2) states that at most p cell towers are to be located. Constraint (3) is just a definitional constraint wherein the fractional coverage of node j at time t by a BS i exists only if BS i is located and node j falls within the coverage area of BS i. Constraint (4) ensures that the number of channels allocated to a demand node is at most equal to its demand at any given time slot. Constraint (5) imposes a restriction on the total channel capacity at any time t. To take into account the fact that coverage of emergency nodes is essential, we have constraint (6), which states that each emergency node should be covered by at least one BS. It should be pointed out here that emergency calls are not considered in assessing the adequacy of channel capacity. This is justified because the volume of such incidents is miniscule in comparison to the overall cellular traffic volume.

There are numerous extensions to this problem that could be possible. For example, while allocating channels, we did not take into account effects such as co-channel interference and adjacent channel interference. We could treat the coverage of the special set of nodes E as a second objective, making it a bicriteria problem. In reality, signal strength varies with distance from the cell tower location. More specifically, under ideal conditions the RSSI value can be expected to decrease inversely in proportion to the square of the distance from the cell tower according to Macario (1997). Furthermore, the effect of foliage, terrain, etc. on RSSI value can be quite significant—see, for example, Delmelle et al. (submitted for publication). According to Akella et al. (2003), calls that have an RSSI value of −89 dB or higher go through uninterrupted (i.e. with probability 1). For values less than −119 dB, the call will not be completed (i.e. with probability 0). When RSSI values fall between −89 and −119 dB, there is a probability associated with call completion. The model can be changed to reflect this by considering partial coverage possibilities of calls. We could use some empirical results to postulate the decrease in signal strength with distance from the cell tower and then develop a partial coverage model to capture this intermediate range of RSSI values. This intermediate range of RSSI values can be particularly relevant since foliage effects can lower RSSI value by as much as 20%, making areas of decent coverage (say with RSSI value of −80 dB) into areas where coverage can be questionable. Thus partial coverage models need to be explored to accurately model the true coverage of both "regular" and "emergency" customers. The NDEC model proposed here is the first of its kind and one of the basic models in cellular network design from the perspective of emergency notification.

3. Model properties

The NDEC formulation is a Mixed Integer Programming (MIP) problem. A closer look at the problem reveals that it is a combination of the set covering and the maximal covering location problems. We could not find any articles that addressed this problem in the OR literature. Church and ReVelle (1974) proposed a maximal covering location problem with mandatory closeness constraints wherein they maximize the population that can be covered within a given service distance S while at the same time ensuring that the users at each point of demand will find a facility no more than T distance away ($T > S$). This is a set covering problem with respect to the distance T and a maximal covering problem with respect to S. The authors solve an example problem but there is no general solution technique proposed in that paper. Since this problem is a superset of the set covering and the maximal covering location problems, it is NP-complete.

The problem becomes more realistic with an increase in the number of time slots. The actual time variation of demand can be modeled accurately only with a large number of time slots but then this presents a very large problem to be solved. For instance, if we take a typical problem with 1000 demand nodes, 200 emergency nodes, 500 potential locations of BSs and 20 time slots, then it would have 10^7 variables and a much larger number of constraints. This presents a very large-scale MIP and professional solvers like CPLEX would not be able to handle such huge data as will be seen later in the paper.

According to the formulation, a demand node that is covered by more than one BS in a solution might not be assigned to the nearest BS. Ideally, a cell phone call made from any location tries to connect to the nearest BS. Arriving at such a solution from a given optimal solution is trivial and there is no change in the objective function value in doing so. In other words, given an optimal solution, it is possible to construct an alternate optimal solution in which each demand node is assigned to its closest BS.

Property 1. *Let $(\boldsymbol{x}^*,\boldsymbol{f}^*)$ be an optimal solution for the problem* **(P1)** *with an objective function value $\boldsymbol{Z}^*$ and let $\boldsymbol{M}^*$ $(\boldsymbol{M}^* \subseteq \boldsymbol{M})$ be the optimal set of BSs. $(\boldsymbol{x}',\boldsymbol{f}')$ is an alternate solution constructed from the original optimal solution such that*

$$x'_i = x^*_i \quad \forall i \in M,$$

$$f'_{ijt} = \begin{cases} \sum_{l\in M^*} f^*_{ljt} & \text{if } i \in M^*,\ A_{ij} = 1 \text{ and } i \text{ is the closest BS to node } j \quad \forall\, j \in N, t, \\ 0 & \text{otherwise.} \end{cases}$$

Then $(\boldsymbol{x}',\boldsymbol{f}')$ is also optimal to the original problem with objective function value $\boldsymbol{Z}^*$.

Proof. First let us prove that the new solution $(\boldsymbol{x}',\boldsymbol{f}')$ is feasible. Since $x' = x^*$, the new set of optimal locations of BSs would be $\boldsymbol{M}^*$. Constraints (2) and (6) are satisfied since $x' = x^*$. Constraint (3) is satisfied for all $f'_{ijt} = 0$. From the definition, $f'_{ijt} > 0$ if $i \in M^*$, $j \in N$ and $A_{ij} = 1$. Hence constraint (3) is satisfied for all i, j, t. If we assume that there is only one BS closest to each node, constraint (4) is automatically satisfied since $\sum_{i\in M} f'_{ijt} = \sum_{i\in M^*} f'_{ijt} = \sum_{i\in M^*} f^*_{ijt} \leqslant 1$. Now,

$$\sum_{i\in M}\sum_{j\in N} H_{jt} f'_{ijt} = \sum_{j\in N} H_{jt} \sum_{i\in M^*} f'_{ijt} = \sum_{j\in N} H_{jt} \sum_{i\in M^*} f^*_{ijt} \leqslant T$$

since f^*_{ijt} is a feasible solution. Hence $\sum_{i\in M}\sum_{j\in N} H_{jt} f'_{ijt} \leqslant T$ which satisfies constraint (5). Therefore, the new solution $(\boldsymbol{x}',\boldsymbol{f}')$ is feasible to the original problem. Now let us compare the objective function values of both the solutions. For the problem **(P1)**,

$$\boldsymbol{Z}' = \sum_{i\in M}\sum_{j\in N} W_t H_{jt} f'_{ijt} = \sum_{j\in N} W_t H_{jt} \sum_{i\in M^*} f'_{ijt} = \sum_{j\in N} W_t H_{jt} \sum_{i\in M^*} f^*_{ijt} = \boldsymbol{Z}^*.$$

Hence $\boldsymbol{Z}' = \boldsymbol{Z}^*$. Therefore, the alternate solution $(\boldsymbol{x}', \boldsymbol{f}')$ is feasible and optimal to NDEC problem. □

We now explore the solution structure to develop heuristics that give a near optimal solution in reasonable time.

Property 2. *Let* $\boldsymbol{F} = \{f_{ijt} | 0 < f_{ijt} < 1, \forall i, j, t\}$. *Then,* $\exists$ *an optimal solution* $(\boldsymbol{x}^*, \boldsymbol{f}^*)$, *such that* $|\boldsymbol{F}| \leqslant$ *number of time slots.*

Proof. Given optimal locations of BSs, for any given time slot, each demand node can be assigned channels equal to its demand until all the channels are used up. In such a case only one demand is satisfied partially. Therefore, we have at most one fractional variable for each time slot and hence the property. □

Property 2 gives insight to the final solution structure. Though the problem is a combination of set covering and maximal covering location problems, it can be modeled as a maximal covering problem by treating the emergency nodes as demand nodes of very high demand, say M (a large number). In such cases, we can use the vast amount of literature available to solve maximal covering location problems to solve this problem. But one intricacy involved would be in cases where the set covering problem is infeasible. The modified problem would not be able to detect any infeasibility in the original problem since the maximal covering location problem is never infeasible. So this kind of approach would work only when the underlying set covering problem is feasible. It should also be noted that the channel capacities should be altered accordingly to accommodate the coverage of emergency nodes of high demand.

4. Solution techniques

4.1. Deterministic Addition (DA) heuristic

The first heuristic considered is called the Deterministic Addition (DA) heuristic. It is similar to the Greedy Addition (GA) Algorithm proposed by Church and ReVelle (1974) for the maximal covering location problem. The idea behind this approach is that, by covering a sparsely covered emergency node, there is a high possibility of covering other emergency nodes that are better covered than this node. Once all of the emergency nodes are covered, the heuristic moves on to the maximal covering problem of the uncovered demand nodes. Now the problem is updated by removing the covered demand and emergency nodes, and by decreasing the total available channels for each time slot by the amount of demand covered in that time slot. Then the DA determines that emergency node with minimum coverage from among the uncovered emergency nodes. Repeating the process again, it selects the BS with maximum weighted coverage covering this emergency node. This process is repeated until all the emergency nodes are covered after which it follows the Greedy Addition (GA) algorithm of Church and ReVelle (1974) to maximally cover the demand nodes. The heuristic terminates if all the demand nodes are covered or p BSs are located or all the available channels are used up. We note that the DA heuristic does not always guarantee a feasible solution even if the original problem is feasible.

4.2. Probabilistic Addition (PA) heuristic

In the PA heuristic there is a probability associated with selecting an emergency node to be covered and this is inversely proportional to the number of BSs covering that emergency node. This heuristic is run for a fixed number of iterations and terminates when it encounters a feasible solution or the iteration limit is reached. This approach is similar to the Simulated Annealing (SA) search technique and helps to prevent the search from getting stuck at local optima. The main advantage of this heuristic is that in most cases it returns a feasible solution (if one exists).

4.3. Set Max Cover (SMC) deterministic heuristic

This is a slight modification of the DA heuristic that concentrates totally on the coverage of emergency nodes in the first phase and then moves on to the coverage of demand nodes in the second phase. This should help remove infeasibilities in the final solution. We call this the SMC heuristic because the first phase of this heuristic involves the set covering problem of the emergency nodes and the second phase involves maximal covering of the demand nodes.

4.4. Set Max Cover (SMC) probabilistic heuristic

This is a modification of the SMC deterministic heuristic with probabilistic selection. This heuristic would try to jump out of local optima (if any) while doing the search.

4.5. Lagrangean heuristic

From our computational experience, the LP relaxation of (**P1**) yielded IP optimal solutions in many cases. But this is not always the case. Motivated by this we develop a Lagrangean heuristic for our problem using subgradient optimization, see e.g., Ravindra et al. (1993). Consider the problem (**P1**). Upon relaxing the constraint set 3 with a penalty cost λ_{ijt} and placing it in the objective function we obtain the formulation

$$L(\lambda) = \text{Maximize} \sum_{i \in M} \sum_{k \in N} \sum_{t} (W_t H_{jt} - \lambda_{ijt}) f_{ijt} + \sum_{i \in M} \sum_{k \in N} \sum_{t} \lambda_{ijt} A_{ij} x_i$$

subject to constraints (2), (4), (5) and (6) and $x_i \in \{0,1\}\ \forall i \in M,\ f_{ijt} \geqslant 0\ \forall i \in M, j \in N, t,\ \lambda_{ijt} \geqslant 0\ \forall i \in M, j \in N, t$.

This problem is separable into two subproblems (**SP1** and **SP2**) in the variables $\boldsymbol{x}$ and $\boldsymbol{f}$:

(**SP1**) Maximize $\displaystyle\sum_{i \in M} \sum_{k \in N} \sum_{t} (W_t H_{jt} - \lambda_{ijt}) f_{ijt}$

subject to

$$\sum_{i \in M} f_{ijt} \leqslant 1 \quad \forall j \in N, t,$$

$$\sum_{i \in M} \sum_{j \in N} H_{jt} f_{ijt} \leqslant T \quad \forall t,$$

$$f_{ijt} \geqslant 0 \quad \forall i \in M, j \in N, t.$$

This problem is further decomposable into separate time slots as

$$(\mathbf{SP1}_t)\quad \text{Maximize} \quad \sum_{i\in M}\sum_{k\in N}(W_t H_{jt} - \lambda_{ijt})f_{ijt}$$

subject to

$$\sum_{i\in M} f_{ijt} \leqslant 1 \quad \forall j \in N,$$
$$\sum_{i\in M}\sum_{j\in N} H_{jt} f_{ijt} \leqslant T,$$
$$f_{ijt} \geqslant 0 \quad \forall i \in M, j \in N.$$

Here, $\mathbf{SP1}_t$ denotes the subproblem **SP1** for time slot t. This is just a linear program in the variable f_{ijt}. Also this resembles a knapsack problem and can be solved without the use of any standard solver. A simple algorithm can be developed to solve the problem to optimality for each time slot. The other subproblem is:

$$(\mathbf{SP2})\quad \text{Maximize} \quad \sum_{i\in M}\sum_{k\in N}\sum_{t}\lambda_{ijt}A_{ij}x_i$$

subject to

$$\sum_{i\in M} x_i \leqslant p,$$
$$\sum_{i\in M} A_{ik}\cdot x_i \geqslant 1 \quad \forall k \in E,$$
$$x_i \in \{0,1\} \quad \forall i \in M.$$

This is a 0–1 IP and is a weighted set covering problem (and hence is NP-complete). Though the problem is NP-complete our conjecture is that, since the number of emergency nodes is very small when compared to the number of demand nodes, finding an IP optimal solution to this problem is relatively easy (for a solver like CPLEX). The rationale behind using Lagrangean relaxation can be summarized as follows:

- One can show through numerical examples that $\mathbf{L}(\lambda)$ does not possess the integrality property, so the best Lagrangean bound will be strictly better than the LP bound.
- For large-scale problems, the number of variables would be of the order of 10^7. Therefore, it would not be sensible to use a solver like **ILOG CPLEX** for solving even LP relaxations.
- The problem decomposes into "manageable" subproblems.
- It generates a lower bound at every iteration since a solution to subproblem **SP2** gives a feasible set of BS locations $\boldsymbol{x}$ to cover emergency nodes. An overall feasible solution can then be generated by inspection.

4.5.1. Subgradient optimization

This technique starts with an initial set of values of the multipliers λ^0_{ijt}. The multipliers are then updated as follows:

$$\lambda^{k+1}_{ijt} = [\lambda^k_{ijt} + \mu_k(f^k_{ijt} - A_{ij}x^k_i)]^+.$$

In this expression, f^k_{ijt} and x^k_i is any solution to the Lagrangean subproblem when $\lambda_{ijt} = \lambda^k_{ijt}$ and μ_k is the step length at the kth iteration. Only the positive part of λ^{k+1}_{ijt} is chosen because they are constrained to be non-negative. To ensure that this method solves the Lagrangean multiplier problem, we need to exercise some care in the choice of the step sizes. If we choose them too small the algorithm would become stuck at the current point and not converge; if we choose the step sizes too large, the iterates λ^k_{ijt} might overshoot the optimal solution and perhaps even oscillate between non-optimal solutions. The following compromise ensures that the algorithm strikes an appropriate balance between these extremes and does converge:

$$\mu_k \to 0 \quad \text{and} \quad \sum_{j=1}^{k} \mu_j \to \infty.$$

For example, choosing $\mu_k = 1/k$ satisfies these conditions. We use the standard subgradient search technique wherein the step size is determined as follows:

$$\mu_k = \frac{\varepsilon_k [L(\lambda^k) - \text{LB}]}{\sum_k \left(f_{ijt}^k - A_{ij} x_i^k\right)^2}.$$

In this expression, LB is a lower bound on the optimal objective function value of the problem (**P1**) and ε_k is a scalar chosen (strictly) between 0 and 2. The denominator is the Euclidean norm of the vector of the inequality constraint that was relaxed. Initially, the lower bound is the objective function value of any known feasible solution to the problem (**P1**). As the algorithm proceeds, if it generates a better feasible solution, it uses the objective function value of this solution in place of the lower bound LB. Since this heuristic has no convenient stopping criteria, we run it for a specified number of iterations and then terminate.

5. Computational results

All five heuristics have been tested on three types of data sets viz. small, medium and large-scale problems. The heuristics have been coded in C, which interfaces with the ILOG CPLEX callable library while solving using the Lagrangean heuristic. These instances have been run on a 768 MB RAM, Intel Pentium 3, 800 MHz processor operating on a Windows platform. The instances for the NDEC problem were created through a small algorithm specially designed for this purpose.

5.1. Instance generator

We developed an algorithm to provide some instances for the NDEC problem. Basically, we had to estimate the size of the working area, the location and number of candidate BSs, the coverage area of each candidate BS, the location and number of demand and crash nodes, the demand of each demand node for each time slot, the total channel capacity and the weight of each time slot.

For all the test problems, a square working area of either 10 units by 10 units or 20 units by 20 units was chosen. The demand and crash nodes were assumed to be randomly distributed over this working area (see Fig. 1). The demands were generated randomly between 1 and 9 (with a mean of 5). This is done over all the

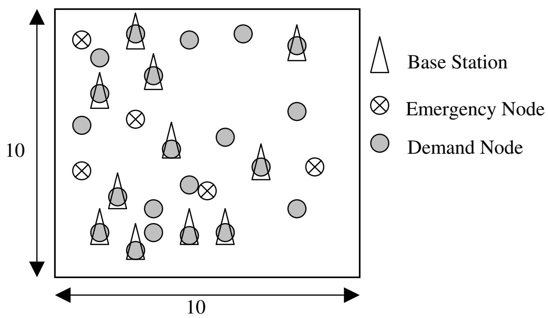

Fig. 1. Instance generator.

time slots. The locations of candidate BSs are assumed to follow a distribution similar to the spatial distribution of the demand nodes. This means that, there would be more candidate BSs where the demand density is high, and less where the demand density is low or 0. Note that the location of a candidate BS coincides with the location of one of the demand nodes. Initially, the BSs are assumed to have a circular coverage area for these test problems. With the coverage obtained, if any candidate BS does not cover a crash node then it is assumed to be covered by the BS nearest to it. This is done to avoid infeasibilities at the initial stages. Hence the coverage region of a BS is not exactly circular. The radius of coverage of a BS is chosen such that, with the actual number of BSs to be located, the whole working area is covered. This was done to avoid any kind of preprocessing to the test problem, which would otherwise reduce the actual number of non-zero variables to a great extent.

Fig. 1 shows a test problem generated using the instance generator. Note that the location of the candidate BSs coincides with the locations of the demand nodes.

5.2. Results

A 10 unit by 10 unit working area was selected to create test problems for the first four heuristics. A total of 300 demand nodes and 50 crash nodes were randomly generated. Based on the spatial distribution of the demand nodes, 200 candidate BS locations were generated. The demand of each node was generated as a random number between 1 and 9 with a mean 5. The total channel capacity was assumed to be 1500. The coverage radius of a BS is 1 unit. The weight of each time slot was assigned on a scale from 1 to 5 randomly. With this input data, the actual number of BSs to be located was varied from 20 (the minimum number required) to 34 and solved using the heuristics. For each value of the actual number of BSs, four different test problems were created and solved using the four separate heuristics. The location of the demand and accident nodes was assumed to be the same for these instances. However, with each new instance we changed the location of the candidate BSs and the demand values. The results are as shown in Table 1.

Table 1 lists the following. Under the CPLEX MIP column, the optimal objective function value and the running time of CPLEX (seconds) is shown. Under each heuristic, the solution is shown as a percentage of the optimal value. In case the heuristic reports an infeasible solution, then the percentage feasibility (i.e. percentage of the crash nodes that it could cover using the assigned number of BSs) is shown in the next column. The third column shows the running time of each heuristic.

Table 2 summarizes the performance of each heuristic. From this table we can conclude that either one of the heuristics or a combination of them can be used to arrive at an effective solution for the NDEC problem. The SMC probabilistic would prove effective only for cases where the actual number of BSs marginally covers the emergency nodes, because it rarely reports infeasibility when the original problem is feasible.

Fig. 2 shows a plot of the feasibility curves for the four heuristics versus the actual number of BSs located. The SMC probabilistic never reports an infeasible solution and hence it shows 100% feasibility in the graph (a parallel line to the X axis). Also note that, as the number of BSs to be located increases, the % feasibility of all the heuristics increases.

5.2.1. Lagrangean heuristic

As mentioned before, the Lagrangean heuristic was tested for three different problem sizes. The results for both the small and medium scale problems are presented in this section. The next section includes a case study that presents the results for large-scale problems.

A test problem on a 10 unit by 10 unit working area was created for the small size problems. The step size (ε_k) was chosen to have an initial value of 2 and if the upper bound did not improve in three successive iterations, the step size was reduced to half. In successive iterations, if the step size reduced to 0.05, it was re-initialized to 2. This was done to avoid the solution from getting stuck at a local value. The instances for

Table 1
Heuristic performance

No. of base stations	CPLEX MIP		Det Addition (DA)			Prob Addition (PA)			Set Max Cover (SMC) det			Set Max Cover (SMC) prob		
	Obj value	Time (seconds)	% optimally	% feasibility	Time (seconds)	% optimality	% feasibility	Time (seconds)	% optimality	% feasibility	Time (seconds)	% optimality	% feasibility	Time (seconds)
20	27,438	132	Inf	88	1	Inf	84	24	Inf	98	1	87.41	100	12
20	27,438	116	Inf	96	1	Inf	76	24	Inf	98	0	78.82	100	12
20	27,438	115	Inf	88	1	Inf	80	23	Inf	94	1	68.78	100	12
20	27,438	116	Inf	92	1	Inf	86	27	Inf	98	1	85.92	100	16
21	27,438	144	Inf	82	1	Inf	80	32	Inf	98	0	58.39	100	17
21	25,571	152	Inf	88	1	Inf	88	26	90.25	100	1	77.45	100	13
21	25,571	115	Inf	90	1	Inf	86	25	Inf	98	1	91.69	100	13
21	30,479	144	Inf	96	1	Inf	82	26	90.04	100	0	77.67	100	13
22	23,937	133	Inf	98	0	Inf	90	26	89.94	100	0	81.88	100	13
22	26,847	133	Inf	96	0	Inf	84	27	83.48	100	0	80.28	100	14
22	21,092	128	Inf	88	1	Inf	90	26	89.63	100	1	81.4	100	14
22	28,746	136	Inf	94	1	Inf	86	26	93.56	100	0	74.86	100	13
23	19,738	136	94.38	100	0	83.47	100	11	88.9	100	1	81.67	100	14
23	28,210	135	Inf	90	1	Inf	86	27	87.08	100	0	79.28	100	14
23	33,362	158	95.33	100	0	88.1	100	12	94.78	100	1	82.26	100	8
23	24,468	164	Inf	98	1	Inf	86	27	87.78	100	1	84.75	100	2
24	31,255	268	99.11	100	1	86.92	100	0	94.25	100	1	85.38	100	2
24	33,835	161	96.42	100	1	83.89	100	5	93.81	100	1	80.1	100	1
24	23,591	144	96.75	100	1	80.52	100	18	87.54	100	1	86.2	100	4
24	30,462	139	Inf	98	0	84.92	100	15	93.24	100	1	81.41	100	1
25	36,095	137	98.23	100	1	85.57	100	4	92.73	100	1	86.92	100	0
25	37,968	178	98.28	100	1	87.47	100	3	96.76	100	1	82.73	100	1
25	27,222	146	Inf	92	1	84.68	100	8	93.92	100	0	74.83	100	4
25	30,400	164	96.84	100	1	90.14	100	2	92.07	100	0	85.05	100	2
26	31,520	241	98.13	100	1	91.64	100	2	96.72	100	1	85.92	100	0
26	35,514	216	96.3	100	1	92.48	100	11	94.04	100	1	80.25	100	1
26	22,173	221	95.93	100	1	89.9	100	1	95.57	100	1	81.21	100	1
26	26,197	172	99.42	100	1	83.29	100	3	96.63	100	1	86.25	100	0
27	30,391	146	98.63	100	1	85.83	100	2	97.74	100	0	88.43	100	1
27	26,107	150	98.08	100	1	83.74	100	1	91.86	100	1	80.48	100	0
27	34,992	157	98.64	100	1	91.36	100	1	95.13	100	1	82.25	100	0
27	28,890	234	98.31	100	1	85.4	100	1	96.99	100	1	83.21	100	0
28	28,441	157	98.61	100	0	88.37	100	1	96.02	100	1	91.3	100	1
28	33,480	138	98.3	100	1	89.27	100	2	97.01	100	1	89.88	100	1

Table 1 (*continued*)

No. of base stations	CPLEX MIP		Det Addition (DA)			Prob Addition (PA)			Set Max Cover (SMC) det			Set Max Cover (SMC) prob		
	Obj value	Time (seconds)	% optimally	% feasibility	Time (seconds)	% optimality	% feasibility	Time (seconds)	% optimality	% feasibility	Time (seconds)	% optimality	% feasibility	Time (seconds)
28	26,878	203	95.7	100	1	87.17	100	1	95.74	100	1	87.91	100	1
28	35,007	161	98.25	100	1	84.12	100	1	97.29	100	1	85.88	100	0
29	31,670	206	98.96	100	1	92.47	100	1	97.34	100	1	87.19	100	1
29	32,017	317	98.95	100	2	83.2	100	1	96.13	100	2	86.45	100	1
29	31,844	279	97.06	100	1	90.39	100	2	88.14	100	1	89.29	100	1
29	34,363	253	98.37	100	1	90.72	100	2	94.42	100	2	83.29	100	1
30	29,617	856	98.75	100	2	83.12	100	3	96.8	100	2	90.94	100	1
30	28,475	282	97.49	100	2	87.3	100	1	97.22	100	1	89.42	100	1
30	31,889	361	97.73	100	2	87.5	100	1	97.61	100	1	87.74	100	2
30	30,170	534	99.27	100	2	84.84	100	1	97.65	100	1	87.07	100	1
31	18,831	332	96.49	100	1	90.81	100	2	96.67	100	2	85.9	100	2
31	38,916	940	99.12	100	2	89.56	100	2	97.66	100	1	92.36	100	2
31	29,834	603	98.24	100	2	92.12	100	1	97.72	100	2	86.24	100	1
31	17,452	745	98.38	100	2	90.41	100	2	97.35	100	1	90.81	100	2
32	27,754	254	98.48	100	1	93.65	100	2	97.67	100	1	94.15	100	2
32	32,189	407	98.49	100	2	92.27	100	1	96.5	100	2	88.37	100	2
32	32,085	239	98.41	100	1	86.54	100	2	97.38	100	2	91.4	100	1
32	27,989	591	98.49	100	2	89.35	100	2	95.29	100	2	84.37	100	2
33	30,591	468	97.07	100	2	89.52	100	2	96.55	100	1	88.13	100	2
33	19,337	260	98.57	100	2	90.34	100	1	96.9	100	2	89.61	100	1
33	27,550	455	98.32	100	1	90.87	100	1	96.12	100	1	89.93	100	1
33	23,418	530	98.57	100	2	92.15	100	1	97.79	100	2	90.51	100	1
34	30,881	280	98.37	100	1	90.57	100	1	97.25	100	1	89.1	100	1
34	25,469	414	99.23	100	2	96.1	100	1	98.2	100	2	93.76	100	1
34	37,185	215	98.39	100	2	90.25	100	1	97.15	100	2	88.04	100	1
34	37,776	507	98.75	100	2	92.13	100	1	97.91	100	1	95.43	100	2

Coverage radius = 1; number of time slots = 10; potential number of BSs = 300; total channel capacity = 1500; number of demand nodes = 300; number of accident nodes = 50.

Table 2
Summary of results for the heuristics

	Average % optimality	Average % feasibility	No. of infeasibilities out of 60	Average computation time
Det Addition (DA)	97.94	97.9	16	1.17
Prob Addition (PA)	88.35	96.4	14	8.44
Set Max Cover (SMC) det	94.62	99.7	6	1.03
Set Max Cover (SMC) prob	84.96	100	0	4.23

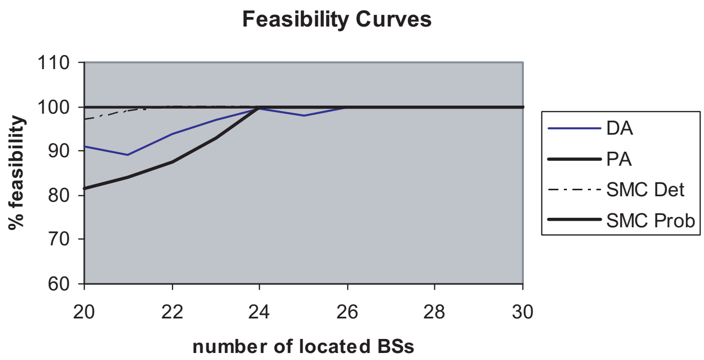

Fig. 2. Feasibility curves for the heuristics.

small size problems consisted of 150 demand nodes, 30 crash nodes, 100 candidate BSs and 10 time slots. The total channel capacity for each time slot was assumed to be 500. The SMC deterministic and probabilistic heuristics provided a starting solution to the Lagrangean heuristic. When varying the actual number of BSs to be located, we create a different instance of the problem but retain the location of the demand and accident nodes. In all the problems the maximum number of Lagrangean iterations was set to 10,000, though this figure was never reached when solving even the large size problems. The subgradient search was terminated under the following criteria:

- starting solution reported by the greedy heuristic is infeasible,
- optimality gap is reduced to less than 1%,
- there is no improvement in the upper bound in 200 successive iterations,
- iteration limit is reached.

The problem is also run on CPLEX 7.1 by creating an LP input file. Note that even in CPLEX the optimality gap was preset to 1%. This means that if the gap between the lower and upper bounds obtained while solving the problem in CPLEX falls to within 1% then the optimization is terminated. Our goal was to test the performance of our algorithm for a preset quality of the solution. The results are shown in Tables 3 and 4. Note that the Lagrangean heuristic performs better in terms of the solution time although in some cases the solution obtained by CPLEX is better than the Lagrangean heuristic. For instance in Table 3, when the number of base stations to be located is 10, the demand weighted coverage obtained by CPLEX is 9860 which is better than the one obtained by the Lagrangean heuristic (9794). With SMC deterministic as the starting solution the first bound obtained by Lagrangean relaxation is good enough to obtain a <1% optimality gap. Hence the solution does not improve from the one provided by the heuristic. In the case of SMC probabilistic, Lagrangean actually builds on the starting solution and finally reaches a <1%

Table 3
Small size problems—SMC deterministic (Lagrangean performance)

No. of base stations	CPLEX		SMC det	Lagrangean heuristic				
	Solution	Time (seconds)	Solution	Lower bound	Upper bound	% optimality gap	Iterations	Time (seconds)
9	14,544	130	14,194	14,467	14,605	0.954	29	6
10	9860	141	9671	9794	9876	0.837	6	2
11	14,724	320	14,605	14,703	14,727	0.163	1	<1
12	13,797	296	13,871	13,871	13,908	0.267	1	<1
13	13,281	305	13,348	13,348	13,400	0.390	1	<1
14	14,125	474	14,125	14,125	14,130	0.035	1	<1
15	13,041	278	13,151	13,151	13,154	0.023	1	<1

Number of demand nodes = 150; number of crash nodes = 30; potential number of base stations = 100; coverage radius = 2; channel capacity = 500; number of time slots = 10.

Table 4
Small size problems—SMC probabilistic (Lagrangean performance)

No. of base stations	CPLEX		SMC prob	Lagrangean heuristic				
	Solution	Time (seconds)	Solution	Lower bound	Upper bound	% optimality gap	Iterations	Time (seconds)
9	12,901	148	11,726	12,901	12,936	0.276	147	42
10	15,178	151	14,579	15,076	15,223	0.975	41	10
11	12,390	250	11,961	12,419	12,452	0.266	24	6
12	15,094	296	14,708	15,071	15,178	0.710	2	<1
13	14,055	337	13,721	14,088	14,139	0.362	1	<1
14	14,886	368	14,898	14,905	14,938	0.221	1	<1
15	14,157	493	13,965	14,121	14,157	0.255	1	<1

Number of demand nodes = 150; number of crash nodes = 30; potential number of base stations = 100; coverage radius = 2; channel capacity = 500; number of time slots = 10.

optimality gap. The Lagrangean heuristic solves the problem in less than 2 seconds average starting with SMC deterministic heuristic whereas CPLEX takes more than 4 minutes. For the SMC probabilistic heuristic, Lagrangean takes approximately 9 seconds whereas CPLEX takes 5 minutes to solve the small size problems. As a conclusion, for small scale problems, the Lagrangean heuristic performs very well both in terms of the quality of the solution and the solution time. CPLEX takes an average of 4 minutes to solve this class of problems.

For medium scale problems, a 20 unit by 20 unit working area was considered. Similar step sizes were chosen. The problem consisted of 500 demand nodes, 150 crash nodes, 400 candidate BSs each with a coverage radius of 3 units and 10 time slots. The total channel capacity for each time slot was assumed to be 2000. The actual number of BSs was varied from 18 to 23.

The SMC deterministic heuristic was used as the starting solution. The Lagrangean heuristic was run with a preset optimality gap of 2% or 10,000 iterations whichever is earlier. Note that in this case the solver CPLEX was run only until the time the Lagrangean heuristic finds a solution. The result is shown in Table 5.

The results look impressive both in terms of arriving at a solution and in terms of computational efficiency. For the CPLEX solver, the problem became intractable. It failed to find a feasible solution within the time in which the heuristic finds a solution with less than 2% gap. A look at the problem tells us that this medium sized problem has approximately 2 million variables and an equal number of constraints. Solving a problem of a million variables is cumbersome even from a point of building input files for the solver. The Lagrangean technique exploits the special structure the problem presents, and reduces the load on CPLEX

Table 5
Medium size problems—SMC deterministic (Lagrangean performance)

No. of base stations	SMC det	Lagrangean heuristic				CPLEX	
	Solution	Lower bound	Upper bound	% optimality gap	Time (seconds)	Solution	Time (seconds)
18	34,592	37,121	37,815	1.87	1000	NA	1000
19	53,522	55,379	56,454	1.94	86	NA	86
20	33,167	34,121	34,659	1.58	79	NA	79
21	45,367	45,491	46,311	1.80	23	NA	23
22	57,643	57,643	58,772	1.96	4	NA	4

Number of demand nodes = 500; number of crash nodes = 150; potential number of base stations = 400; coverage radius = 3; total channel capacity = 2000; number of time slots = 10.

to solve a 0–1 IP of just 400 variables (at each iteration). The knapsack problem is solved without using CPLEX by a simple algorithm. This involves building huge arrays and running search algorithms on these arrays. But comparatively this is much easier to do than to write an LP format file to CPLEX. Though the Lagrangean heuristic technique exposes some of the limitations of professional solvers, at the same time it advocates clever usage of solvers by solving tractable problems iteratively and searching for the best solution. In the results above, an interesting feature to be observed is that the solution time of the heuristic decreases with an increasing number of BSs, whereas the CPLEX solution time exhibits no such behavior. The main reason behind this is that the starting solution provided by the greedy heuristic improves with more BSs as the problem moves away from infeasibility. Since the initial solution quality is better, improving this further using the Lagrangean approach (to within 2% optimality) would consequently require lesser time.

From our computational experience, we observed that given a small-sized problem where the radius of coverage is smaller for each BS, CPLEX solves it as efficiently as the Lagrangean heuristic. This is because of the fact that, with a weak coverage of each BS, the number of non-zero variables is greatly reduced and the resulting problem size becomes very small. But we focused on testing the performance of our heuristic on a real world problem where even after preprocessing, the resulting number of non-zero variables is very large. Moreover, we have not included any preprocessing techniques in the heuristic. And owing to this, when the subgradient search is done, the program scans the entire variable set even though preprocessing can ignore some of them. This, if implemented this would actually lead to a considerable reduction in the problem size and subsequently the solution time.

6. Case study

We have applied the Lagrangean heuristic technique to the rural parts of Erie County, New York, where the probability of automobile crashes is known. The motivation behind choosing this as our study region is twofold. First, the Emergency Medical Services (EMS) response time in rural areas is much greater than in urban areas, and it is precisely these areas of the County in which ACN offers the greatest promise. Secondly, for the consistency of our study, we used the same rural crash data as the one discussed in Akella et al. (2003).

6.1. Background information

Rural areas of Erie County represent villages and towns excluding the City of Buffalo and its immediate suburbs. They cover approximately 61% of the county. The rural areas are slightly hilly, especially in

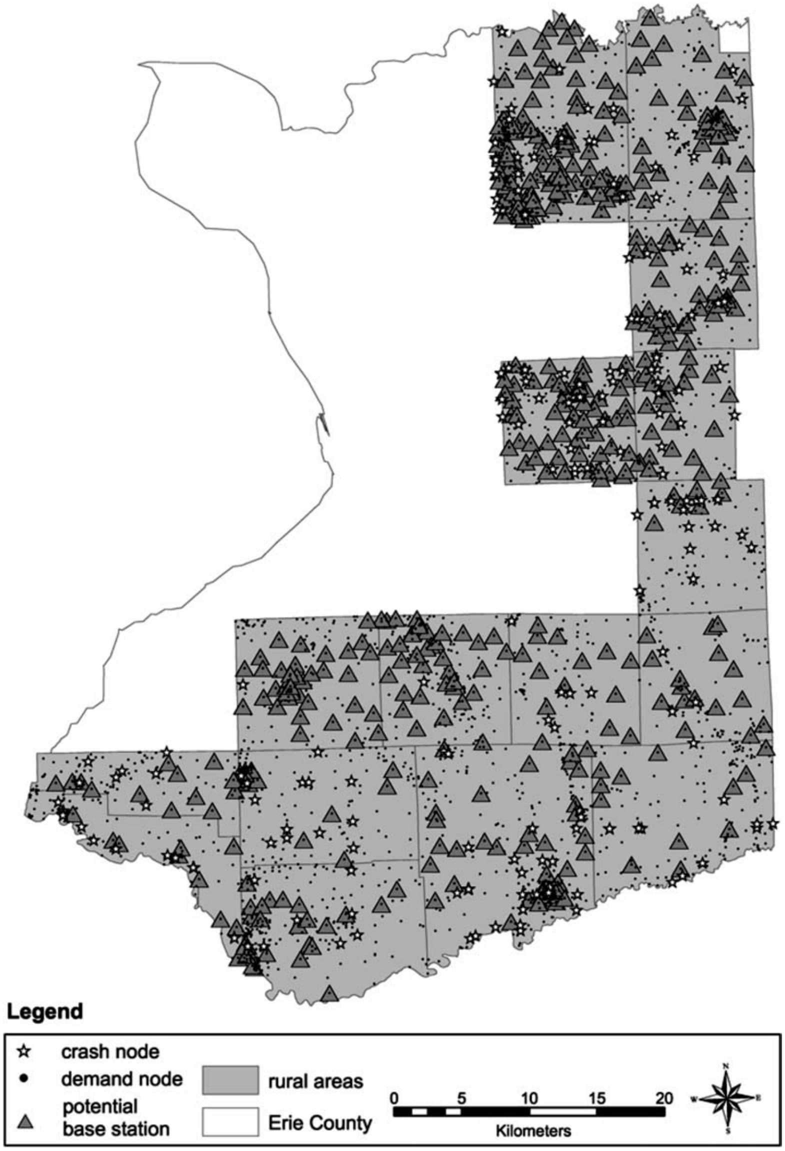

Fig. 3. Location of crash, demand nodes and potential base stations in rural Erie County, New York.

the southeastern corner. The population density in 2000 was 72 people per square kilometer (NYSDOT, 2000).

6.2. Data

For emergency nodes, the location of the 210 rural crashes of 1995 discussed in Akella et al. (2003) was used. The demand nodes were represented using the centroid of rural census blocks. A census block is a small area bounded by a series of street, roads, railroads, streams or any visible features. Census blocks are the smallest geographic areas for which the Census Bureau collects and tabulates decennial census data. The total population of each census block (Census, 2000) was assigned to its centroid. There are 2336 census block centroids in rural Erie County. A total of 1824 blocks were used since some do not contain population information. Fig. 3 shows the distribution of the demand and crash nodes in rural Erie County. Five hundred candidate BS locations were chosen from among the 1824 demand nodes in the region. The choice was made based on the spatial distribution of demand. The coverage radius of each BS is assumed to be 3 km. However, in case a crash node is not covered by any of the candidate BSs, it is then assumed to be covered by the BS closest to it. The radius of coverage has been chosen to be 3 km based on the results of

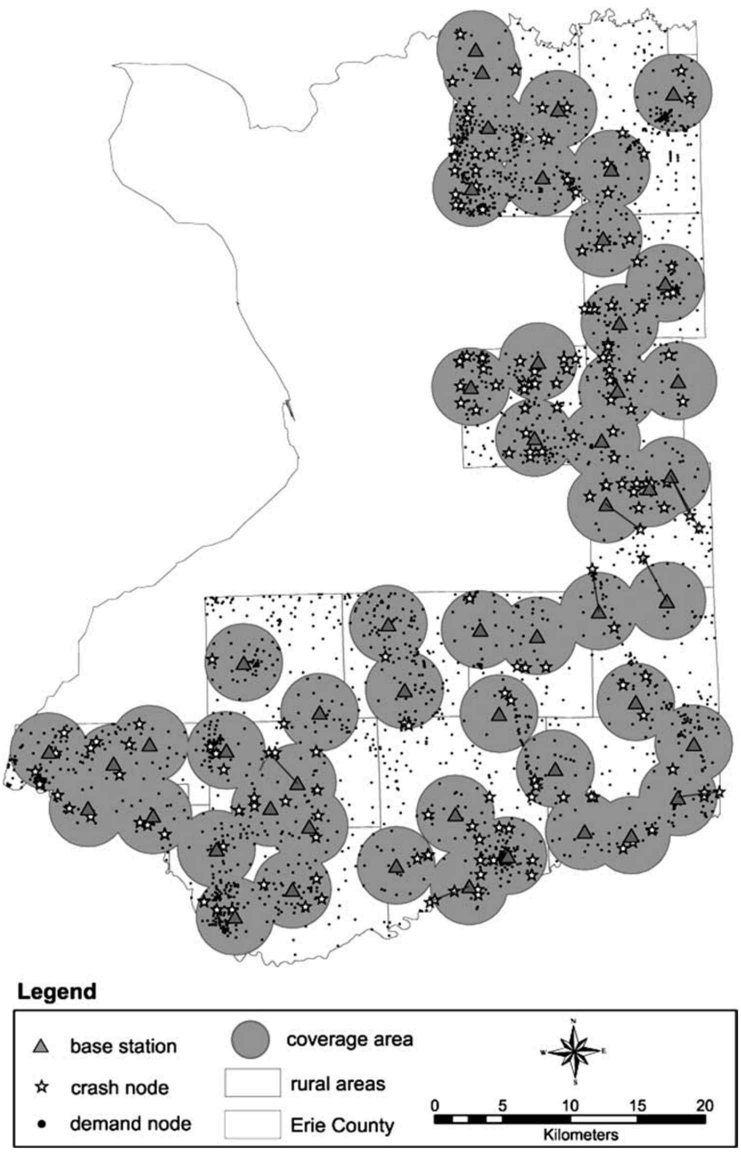

Fig. 4. Optimal base station locations in rural Erie County (3% optimality gap).

Table 6
Case study results (Lagrangean performance)

No. of base stations	CPLEX		Lagrangean heuristic		
	Solution	Time (seconds)	Average % optimality gap (preset)	Average iterations	Average time (seconds)
51	NA	NA	4.973(5)	373.75	3983.50
51	NA	NA	2.829(3)	718	7194.75
51	NA	NA	1.834(2)	1077.50	16,947.75

Number of demand nodes = 1824; number of crash nodes = 210; potential number of base stations = 500; coverage radius = 3; channel capacity = 80,000; number of time slots = 10.

experiments in Delmelle et al. (submitted for publication). They empirically tested the decay of RSSI value with distance from a base station. From their results it was found that a 3 km distance yields a strong signal that essentially guarantees call completion (assuming the availability of a channel). In summary, our problem is to find 51 optimal BS locations that cover the crash nodes and maximally cover the demand nodes.

6.3. Results

The performance of the Lagrangean heuristic is shown in Table 6. The problem was solved for three different instances each with a preset optimality gap of 5%, 3% and 2% respectively. For each value of the preset optimality gap, we ran four different instances and then report the average performance. This means that the heuristic would terminate if it finds a feasible solution within the preset optimality gap. The instances are different in terms of the locations of the candidate BSs and the values of the demands over all the time slots. Note that the same problem has not been solved for each instance. An entirely different problem was created and solved. This has been done to ensure the performance of the heuristic over a range of problems and simultaneously to study the solution time with increasing complexity. NA under the CPLEX column indicates that the solver was unable to find a feasible solution. In fact for these problem instances, the solver fails to read the input format files. The problem size is roughly 10 million variables and an equal number of constraints.

From our computational experience for this class of problems, we observed that the Lagrangean starts with an initial optimality gap of around 35% (this is the gap between the starting solution and the first upper bound obtained) and improves it to less than 5%. This is a remarkable improvement in the gap though the actual solution improves by around 10%. We have the liberty of stopping the solution at any preset gap to work a tradeoff with solution time. The solution time increases greatly as the optimality gap is reduced. As shown in the table, with a preset gap of 2% the heuristic takes nearly 5 hours 30 minutes to solve the problem. But this is reasonable considering the fact that the solver fails to read the problem. Fig. 5 shows the subgradient search behavior for a solution with 5% optimality gap. The solution starts with a large initial gap between the upper and lower bounds. The gap later reduces to less than 5%. This figure visually demonstrates the classical subgradient optimization search technique.

The optimal location of the Base Stations and the coverage obtained by the solution is shown in Fig. 4. Note that there is a link connecting uncovered crash nodes to BSs. This should be interpreted as that crash node being covered by a BS to which it is connected by a link. As explained before, this was a measure taken to avoid unwanted infeasibilities in the problem. As is seen from the figure, we have BSs located in those regions where the demand density is high and where there are crash nodes. The coverage area shown in the figure is a buffer drawn around a BS with 3 km radius. This would be the

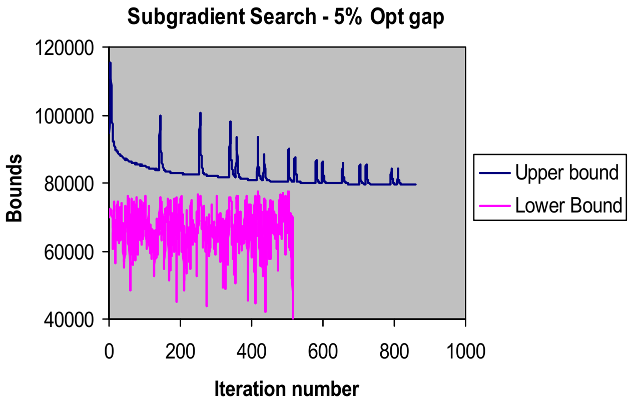

Fig. 5. An example of subgradient search for case study—5% gap.

coverage of a BS with respect to the demand nodes. There are a lot of uncovered demand nodes in the final solution. This is due to the fact that the number of BSs was limited to 51. Increasing this number would give better coverage with respect to the demand nodes while maintaining the coverage of the crash nodes.

7. Conclusions and future research

A cellular network design problem has been addressed from the perspective of emergency notification. The problem has been formulated as a Mixed Integer Program (MIP). Several properties that help in gaining a deeper insight into the problem structure have been developed. Four different solution techniques have been proposed that produce high quality solutions in reasonable time. Finally, a Lagrangean heuristic is developed that takes a starting solution from one of the above heuristics and performs a subgradient search to improve the optimality gap. Results show that the Lagrangean heuristic performs remarkably well when compared to the ILOG CPLEX solver for all kinds of problem sizes. Finally, a case study is presented that applies this solution technique to a practical problem in the rural sections of Erie County, New York.

The Lagrangean heuristic technique developed can be further improved in terms of its computation time by exploiting the LP nature of the knapsack problem. Instead of solving it as a knapsack every time the multiplier is updated, one can input the problem to CPLEX after some preprocessing. The main reason behind this is that at successive iterations, we are effectively solving the same problem with some changed coefficients. With the use of certain features embedded in CPLEX to re-optimize a LP problem, the computational burden of re-solving it from the beginning may be reduced. Also, efficient preprocessing techniques can be incorporated in the heuristic to eliminate redundant variables and constraints. This is a suggested future research direction.

Though we assumed that the signal strength varies deterministically, in reality it does not do so. There would be a probability associated with covering a customer at a given point. A stochastic model that incorporates this feature and maximizes the expected coverage would come closer to real world problems in rural areas, where emergency coverage is important. This is another suggested future research direction.

Acknowledgements

This project has been funded by a grant from the Center for Transportation Injury Research (CenTir) to the University at Buffalo. This support is gratefully acknowledged. We would also like to thank the referees for their valuable comments on an earlier version of this paper.

References

Akella, M.R., Bang, C., Beutner, R., Delmelle, E.M., Wilson, G., Blatt, A., Batta, R., Rogerson, P.A., 2003. Evaluating the reliability of automated collision notification (ACN) systems. Accident Analysis and Prevention 35, 349–360.

Blow, O., Magliore, L., Claridge, J.A., Butler, K., Young, J.S., 1999. The golden hour and the silver day: Detection and correction of occult hypo perfusion within 24 h improves outcome from major trauma. Journal of Trauma, Injury, Infection, and Critical Care 7 (5), 964–969.

Bose, R., 2001. A smart technique for determining base-station locations in an urban environment. IEEE Transactions on Vehicular Technology 50 (1), 43–47.

Chamaret, B., Josselin, S., Kuonen, P., Pizarroso, M., SalasManzanedo, N., Wagner, D., 1997. Radio network optimization with maximum independent set search. In: Proceedings of the IEEE/VTS 47th Vehicular Technology Conference, Phoenix, USA, vol. 2, 1997, pp. 770–774.

Champion, H.R., Augenstein, J.S., Cushing, B., Digges, K.H., Hunt, R., Larkin, R., Malliaris, A.C., Sacco, W.J., Siegel, J.H., 1998. Automatic crash notification: The public safety component of the intelligent transportation system. AirMed, p. 36.

Church, R.L., ReVelle, C., 1974. The maximal covering location problem. Regional Science 30, 101–118.

Clark, D.E., Cushing, B.E., 2002. Predicted effect of crash notification on traffic mortality. Accident Analysis and Prevention 34, 507–513.

Daskin, M.S., 1995. Network and Discrete Location: Models, Algorithms and Applications. John Wiley and Sons, Inc., New York.

Delmelle, E., Rogerson, P., Akella, M., Batta, R., Blatt, A., Wilson, G. A spatial model of remote signal strength indicator values. Transportation Research C, submitted for publication.

Evanco, W.M., 1999. The potential impact of rural mayday systems on vehicular crash fatalities. Accident Analysis and Prevention 31, 455–462.

Federal Communications Commission (FCC), 1996. Revision of the Commission's Rules to Ensure Compatibility with Enhanced 911 Emergency Calling Systems. Available from <http://www.nena.org/govtaffairs/Leg/94-102%20ORIGINAL.pdf>.

Fritsch, T., Tutschku, K., Leibnitz, K., 1995. Field strength prediction by ray-tracing for adaptive base station positioning in mobile communication networks. In: Proceedings of the 2nd ITG Conference on Mobile Communication'95, Neu Ulm.

Ibbetson, L.J., Lopes, L.B., 1997. An automatic base site placement algorithm. In: Proceedings of the IEEE/VTS 47th Vehicular Technology Conference, Phoenix, USA, vol. 2, pp. 760–764.

Jacobs, L.M., Sinclair, A., Beiser, A., D'Agostino, R.B., 1984. Prehospital advanced life support: Benefits in trauma. Journal of Trauma 24, 8–13.

Lerner, E.B., Moscati, R.M., 2001. The golden hour: Scientific fact or medical "Urban Legend"? Academic Emergency Medicine 8, 758–760.

Macario, R.C.V., 1997. Cellular radio: Principles and design. McGraw-Hill Series on Telecommunications, second ed., p. 276.

National Highway Transportation & Safety Authority (NHTSA), 2000. Traffic Safety Facts 1999. Overview, DOT HS 809 092.

New York State Department of Transportation, 2000. 2000 highway mileage report. Highway Data Service Bureau. Available from <http://www.dot.state.ny.us/tech_serv/high/files/erie.pdf>.

Ravindra, A.K., Thomas, M.L., James, O.B., 1993. Network Flows—Theory, Algorithms and Applications. Prentice Hall.

Sherali, H.D., Pendyala, C.M., Rappaport, T.S., 1996. Optimal location of transmitters for micro-cellular radio communication system design. IEEE Journal on Selected Areas in Communications 14 (4), 662–673.

SmartRisk, 1998. The economic burden of unintentional injury in Canada.

Statamatelos, D., Ephremides, A., 1996. Spectral efficiency and optimal base placement for indoor wireless networks. IEEE JSAC 1, 434–438.

Tcha, D.W., Myung, Y.S., Kwon, J.H., 2000. Base station location in a cellular CDMA system. Telecommunication Systems 14, 163–173.

Tutschku, K., Gerlich, N., Tran-Gia, P., 1996. An integrated approach to cellular network planning. In: Proceedings of the 7th International Network Planning Symposium (Networks 96), pp. 185–190.

U.S. Bureau of the Census, 2000. Available from <www.geographynetwork.com/data/tiger2000/index.html>.

Walker, J., 1996. Methodology application: Logistics regression using the crash outcome data evaluation system (CODES) data. Technical report for the National Highway Traffic Safety Administration (NHTSA) and National Center for Statistics and Analysis (NCSA) in Department of Transportation, September.

Wright, M.H., 1998. Optimization methods for base station placement in wireless applications. In: IEEE Vehicular Technology Conference Proceedings, pp. 387–391.