Survivability and performance optimization of mobile wireless communication networks in the event of base station failure

Kuo-Chung Chu [a,*], Frank Yeong-Sung Lin [b,1]

[a] *Department of Information Management, Jin-Wen Institute of Technology, No. 99, An-Chung Road, Shin-Tien City, Taipei County 231, Taiwan*
[b] *Department of Information Management, National Taiwan University, No. 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan*

Abstract

In this paper, we investigate the survivability of mobile wireless communication networks in the event of base station (BS) failure. A survivable network is modeled as a mathematical optimization problem in which the objective is to minimize the total amount of blocked traffic. We apply Lagrangean relaxation as a solution approach and analyze the experiment results in terms of the blocking rate, service rate, and CPU time. The results show that the total call blocking rate (CBR) is much less sensitive to the call blocking probability (CBP) threshold of each BS when the load is light, rather than heavy; therefore, the more traffic loaded, the less the service rate will vary. BS recovery is much more important when the network load is light. However, the BS recovery ratio (BSRR), which is a key factor in reducing the blocking rate for a small number of BSs, is more important when a system is heavily loaded. The proposed model provides network survivability subject to available resources. The model also fits capacity expansion requirements by locating mobile/portable BSs in the places they are most needed.
© 2006 Elsevier Ltd. All rights reserved.

Keywords: Base station recovery; Lagrangean relaxation; Mathematical modeling; Network survivability; Performance evaluation; Quality of service

1. Introduction

Although the survivability of telecommunication networks has been thoroughly investigated, it is still a crucial issue for both service providers and end users. Service providers increase competition by provisioning and deploying reliable/survivable network services. Actually, the problem of network survivability stems from network planning. The planning process can be divided into the following phases: topological design, traffic routing/management, and circuit routing design of the transmission facility network [1]. In traffic management, the

---

* Corresponding author. Tel.: +886 2 8212 2323; fax: +886 2 8212 2339.
  E-mail address: kcchu8992@gmail.com (K.-C. Chu).
[1] Tel.: +882 2 3366 1191.

primary goal is to provide optimal capacity and meet quality of service (QoS) requirements, while minimum-cost routing is primarily concerned with circuit routing design.

Nowadays, a fiber optic-based backbone is incorporated in the design of network topologies. As a single transmission link carries a significant amount of traffic, failure of such a link causes a breakdown of services. Failures can occur as a result of natural disasters (floods, hurricanes, etc.), human actions (war, terrorism, e.g., 911), or failure of software or control systems. Thus, the network must be designed for survivability so that traffic can still be carried immediately after a failure. Efforts should focus on designing networks/protocols in such way that service can be maintained at a reasonable cost if there is a failure. To improve the performance of the network in such an event, the following network planning and management implementation options should be considered: dynamic call routing between switching nodes, circuit diversity through cross-connect nodes, augmented trunk capacity, real-time restoration in the transmission facility network, or a combination of these options [2–4]. Finding an approach that meets survivability requirements is generally regarded as an NP-hard problem [5].

To date, most research has concentrated on survivability in wired networks. In general, a wireless or mobile network consists of a number of components, including base stations (BSs), BS controllers (BSCs), a mobile switch center (MSC), home location registers (HLRs), visiting location registers (VLRs), signaling system 7 (SS7), and high-capacity trunks. As a failure could involve one or more of these components, the survivability of mobile wireless networks is also an important research area [6,7]. The impact of a failure can be measured in terms of the number of users affected and the duration of the outage.

Table 1 summarizes the effects of failure and the mitigation strategies in a wireless environment [6]. A BS serves hundreds of mobile users (MUs) who either initiate new calls or keep their on-going (handoff) calls, while a BSC serves thousands of MUs by supporting switching between several neighboring BSs. An MSC, which is a switch that interconnects a number of BSCs, is capable of supporting at least 100,000 MUs. The HLR is a central database that contains details of each MU authorized to use the network. The VLR is a temporary database of MUs that have roaming in the particular area the database serves. Each BS in the network is served by exactly one VLR; hence, an MU cannot be present in more than one VLR at a time. The failure of switch level components significantly affects services, so that most MUs are unable to access service due to the loss of initial call requests or the end of call delivery. Fortunately, by using redundant components the switches can be operated reliably, and in the event of failure of HLRs/VLRs, the replicated databases are sufficient to continue service.

Even though redundant components can be used to mitigate the impact of a BS failure, the cost is an important issue. In addition, wireless network survivability has to take account of radio channel capacity, interference, and user mobility; as well as the impact a failure has on the signaling network and how a failure in one part of a network impacts on several other parts of the network. To provide reliable and survivable wireless and mobile services, network providers must find ways to reduce the number of network failures and cope with failures when they do occur.

The remainder of this paper is organized as follows. Section 2 contains a literature review. In Section 3, a mathematical problem formulation of a survivable network is proposed. Section 4 presents a solution approach to the problem based on Lagrangean relaxation. Section 5 describes the computational experiments. In Section 6, a sensitivity analysis of associated constraints is conducted by calculating the Lagrangean multipliers for all procedures. Finally, in Section 7, we present our conclusions.

Table 1
The effects of wireless component failure and mitigation strategies

Failure	Cause	Number of users affected	Time to fix	Mitigation strategies
BS/BSC	Hardware, software, nature	1000–20,000	Hours to days	Overlay BS, redundant components
MSC	Hardware, software, operators	100,000	Hours to days	Spare components, power, smaller switches
HLR/VLR	Hardware, software	100,000	Hours to days	Replicated database, redundant components

2. Literature review

The next-generation mobile environment of CDMA-based (code division multiple access-based) networks will have a variety of requirements, such as multimedia data services, higher data rates, mobility, and QoS [8]. With regard to QoS, several previous works have addressed the issue of multimedia traffic in terms of call admission control (CAC). Generally, a network system manages available resources and allocates them in an optimal way among the system's users. Survivable service is another important issue in wireline/wireless communications networks [2–6].

A typical infrastructure of a cellular network consists of a number of components, as shown in Fig. 1. A failure in an MSC affects nearly all customers covered by it – perhaps hundreds of thousands of users. Even though failed components can be replaced with spare parts, natural disasters, such as floods and earthquakes, as well as some human factors, may cause a BS to crash. In a wireless system, a BS that communicates directly with MUs is a critical facility. If more BSs are deployed, the coverage as well as the system's capacity can be enhanced, but there is a greater possibility of BS failure. Thus, partial users may be out-of-service and the overall QoS would be degraded.

In terms of operating, if we can properly recover some of the failed BSs, it would enhance QoS and provide survivable service. Accordingly, BS recovery is one of the most important approaches for minimizing the total system call blocking rate (CBR). Intensive research on the comparison of FH-CDMA and DS-CDMA for wireless survivable networks was reported in [6], but it did not deal with the overall call blocking problem in conjunction with BS recovery. Even though [9] dealt with the survivability issue as an integrated planning problem, it focused on the long-term planning problem; however, network monitoring/servicing to identify and avoid potential failures in a system's operation is more important in the short-term.

The basic model of network survivability proposed in [10] uses an exhaustive search approach. It denotes a set of failed BSs and a set of fixed BSs as $|F|$ and $|F'|$, respectively. The overall procedure of the basic model is as follows:

Step 1. Specify the number of BSs $|F'|$ to be repaired, after which the specific number of combinations required to fix the BSs can be generated, say $C(|F|, |F'|)$.
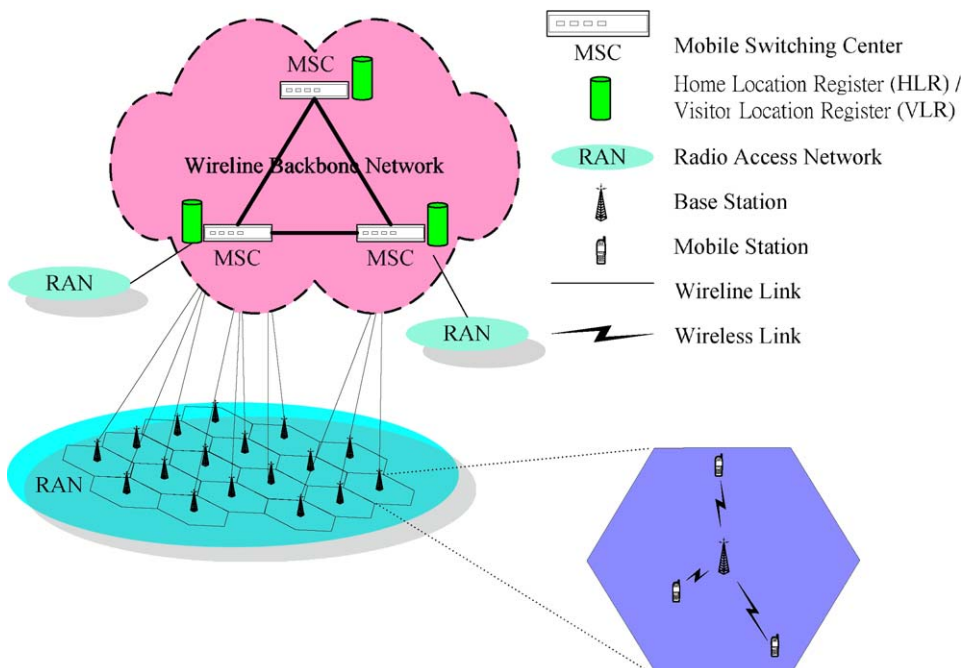

Fig. 1. System architecture of integrated wireline/wireless networks.

Table 2
Definition of sets for network survivability

Notation	Description
B	A set of BSs
B'	A set of available BSs, $B' = B - F$
B''	A set of workable BSs, $B'' = B' \cup F'$
F	A set of failed BSs, $F \subset B$
F'	A set of fixed BSs, $F' \subset F$
T	A set of MUs

Step 2. Sequentially repair the BSs selected from each combination identified in Step 1.

Step 3. Solve the network survivability problem of all workable BSs and calculate the minimal CBR value of each combination; compare the optimal values of all combinations and retain the overall optimal value; if end of total combinations, go to Step 4, or Step 2 otherwise.

Step 4. End solution procedure.

The time complexity of the basic model is $O(|F|^2)$. However, in this paper, we propose a combinatorial optimization model (hereafter called the extended model) that reduces the complexity to $O(1)$. We focus on BS recovery decisions to formulate the survivability problem in terms of call blocking control. The objective is to minimize the blocked traffic in the overall system by allocating some restricted resources to recover broken BSs. To better describe the problem, the sets for modeling the problem are listed in the Table 2. We assume there are $|B|$ BSs in a system, and that a number of BSs, say $|F|$, have failed. To minimize the total blocked traffic, some BSs, say $|F'|$, can be fixed by allocating restricted resources. The BS recovery ratio (BSRR) is defined as $|F'|/|F|$. A system that enables available BSs (i.e., functioning) to cooperate with fixed BSs provides survivable service with overall minimal traffic blocking.

3. Network survivability and performance modeling

3.1. Problem description

The main advantage of the extended model is the time complexity, which is reduced from $O(|F|^2)$ to $O(1)$, where $|F|$ is the total number of failed BSs. In the basic model [10], the decision of each BSRR is selected from the minimal objective values among all recovery combinations. For example, if $|F| = 4$ and BSRR = 0.5 (i.e., the number of BSs to be fixed is two), there exist $C(4, 2) = 6$ possible cases. However, instead of using the time-consuming approach in the basic model, the extended model considers the BS recovery decision for network survivability and performance modeling with CBR minimization jointly. The more traffic loaded into the system, the greater the improvement in time consumption. To effectively evaluate the performance of network survivability, the model only focuses on voice service. We do not consider the existing MU connections,
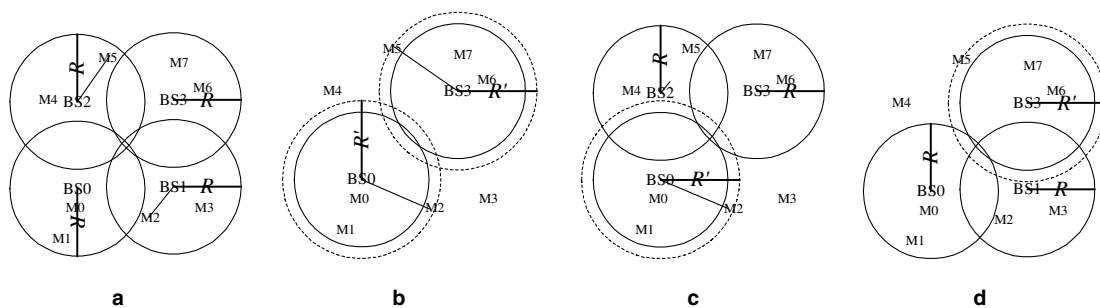

Fig. 2. An example of network survivability: (a) normal case; (b) BS1 and BS2 failures with BS0 and BS3 power adjustment; (c) BS2 recovery with BS0 power adjustment; and (d) BS1 recovery with BS3 power adjustment.

but we do assume perfect power control, the uplink perfectly separated from the downlink, fading, and down-link analysis. In order to simplify the analysis, some complicated scenarios like new, re-homing, outbound, and handover calls are not dealt with.

The extended model provides network survivability by allocating redundant resources. Furthermore, over-laying the architecture, by adjusting the transmission power of all available/workable BSs, reduces the impact of BS failures. Fig. 2 shows an example of four BSs (BS0–BS3) and eight MUs (M0–M7), where two MUs are covered by each BS. In Fig. 2(a), all BSs work normally (no failures) within power radius R. M2 and M5 are served by BS1 and BS2, respectively. In Fig. 2(b), both BS1 and BS2 are failures. Thus, both BS0 and BS3 adjust power within R' to serve M2 and M5, respectively. However, it is a complicated process to recover some failed BSs. Fig. 2(c) and Fig. 2(d) show alternative ways of recovering BS2 and BS1, respectively. The power of the available BSs, BS0 and BS3, is also adjusted. Network survivability deals with recovery decisions about failed BSs, i.e., recovering N of M failed BSs $C(M, N)$, is a combinatorial optimization problem.

3.2. Performance indicators

Snow et al. [6], in their study of the performance issues of wireless communications, propose an outage index that considers the magnitude (the number of customers affected) and duration (the length of time that service is disrupted). In addition, the index focuses on two service components: (1) registration blocking (I_{RB}), i.e., new customers are unable to register with the wireless network; and (2) call blocking (I_{CB}), i.e., the failure of calls by registered customers. Based on I_{RB} and I_{CB}, we now introduce two performance indicators to be analyzed in this research, namely, the service rate and the call blocking rate.

(1) Service rate (R_S): No matter what causes BSs to fail, the power of available/workable BSs (integrating partially fixed BSs) is adjusted. Some MUs originally covered by the failed BSs will probably still be out of service, which will result in registration blocking, $R_S = (1 - I_{RB})$.

(2) Call blocking rate (R_B): The call blocking probability (P_B) of MUs is expressed by Erlang's B formula $P_B = B(g, c)$, where g is the aggregate traffic, and c denotes the available channel resources. Thus, CBR is $R_B = g \cdot P_B = g \cdot B(g, c)$. Actually, in a survivable environment, P_B can be further decomposed into two cases, (i) $P_B = B(g, c)$, where g is aggregate flow of total users covered by workable BSs, and c denotes all available channels; and (ii) $P_B = B(g', c') = 1$, where g' is the aggregate flow of lost traffic (out of service); and c' is zero, since the system cannot allocate any channels.

3.3. Performance modeling

The extended model is formulated as the following mathematical optimization problem (IP). The objective is to minimize the total CBR of the overall system, which is the sum of the average CBR traffic in the system and lost traffic (g'). Associated parameters used in the model are defined in Table 3. The decision variables (DVs) of the extended model are defined in Table 4

$$Z_{IP} = \min \sum_{j \in B''} g_j BB_j(g_j, c_j) + g' \tag{IP}$$

subject to:

$$\left(\frac{E_b}{N_{total}}\right)_{req} \leqslant \frac{\frac{P}{N_0}}{\left(1 + \frac{1}{G}\alpha\frac{P}{N_0}(c_j - 1) + \frac{1}{G}\alpha\frac{P}{N_0} \sum_{\substack{j' \in B'' \\ j' \neq j}} \left(\frac{\frac{r_{j'}}{2}}{\text{Max}\left(D_{jj} - \frac{r_{j'}}{2}, \varpi\right)}\right)^{\tau} c_{j'}\right)} \quad \forall j \in B'' \tag{1}$$

$$\sum_{t \in T} A_t z_{jt} = g_j \quad \forall j \in B'' \tag{2}$$

$$D_{jt} z_{jt} \leqslant r_j u_{jt} \quad \forall j \in B'', \ t \in T \tag{3}$$

Table 3
Definitions of parameters in the extended survivability model

Notation	Description
ϖ	A small number
α	Voice activity factor
τ	Attenuation factor
β_j	Threshold of the call blocking probability (CBP) for each BS j, $j \in B$
A_t	The traffic requirement of MU t (in Erlangs), $t \in T$
b'	The artificial BS that carries a rejected call when the admission control function decides to reject the call request
$BB_j(g_j, c_j)$	CBP of the Erlang-B function with aggregate traffic g_j and available channel c_j in BS j
D_{jj}	Distance between base station j and j'
D_{jt}	Distance between BS j and MU t
E_b	The energy that BSs receive
G	The processing gain
M_j	Upper bound (UB) on the number of users that can be active at the same time in BS j, $j \in B$
N_0	Background noise
N_{Total}	Total noise
P	The power a BS receives from an MU that is homed to the BS with perfect power control
R_j	Upper bound on the transmission power radius of BS j, $j \in B$
s_j	Indicator function of BS j, which is 1 if BS has failed, and 0 otherwise
U	UB on number of failed BSs to be repaired
u_{jt}	The coverage indicator function, which is 1 if MU t can be served by BS j, and 0 otherwise
Y_j	The set of transmission radii of BS j

Table 4
Decision variables in the extended survivability model

Notation	Description
c_j	The number of MUs active at the same time in BS j
f_j	DV, which is 1 if failed BS j is fixed, and 0 otherwise
$f_j \oplus s_j$	DV, which is exclusive-or (XOR) function of f_j and s_j, and expressed by $(1 - f_j)(1 - s_j)$
g'	DV of aggregate lost traffic in the system, where $g' = \sum_{t \in T} A_t \sum_{j \in B''}(1 - z_{jt})$
g_j	DV of aggregate traffic in BS j
r_j	DV of transmission power radius in BS j
z_{jt}	DV, which is 1 if MU t is served by BS j, and 0 otherwise

$$\sum_{j \in B'' \cup b'} z_{jt} = 1 \quad \forall t \in T \tag{4}$$

$$u_{jt} \leqslant f_j \oplus s_j \quad \forall j \in B, \ t \in T \tag{5}$$

$$f_j \leqslant s_j \quad \forall j \in B \tag{6}$$

$$\sum_{j \in F} f_j \leqslant U \tag{7}$$

$$c_j \leqslant M_j \quad \forall j \in B'' \tag{8}$$

$$0 \leqslant r_j \leqslant R_j \quad \forall r_j \in Y_j, \ j \in B'' \tag{9}$$

$$BB_j(g_j, c_j) \leqslant \beta_j \quad \forall j \in B'' \tag{10}$$

$$z_{jt} = 0 \text{ or } 1 \quad \forall j \in B'', \ t \in T \tag{11}$$

$$f_j = 0 \text{ or } 1 \quad \forall j \in B \tag{12}$$

$$c_j \in Z^+ \quad \forall j \in B'' \tag{13}$$

Constraint (1) ensures that each traffic demand is serviced by a BS with the required QoS. The second and third terms of the denominator represent intra-cell and inter-cell interference, respectively. Constraint (2) calculates

the aggregate flow of each BS j. Constraint (3) requires that an MU must be in the coverage area of a BS before it can be served by that BS. Constraint (4) guarantees that each MU cannot home into more than one BS. Constraint (5) ensures that $u_{jt} = 1$ iff $f_j \oplus s_j = 1$ in each of the following cases: (i) BS j has failed ($s_j = 1$) and has to be fixed ($f_j = 1$); and (ii) BS j is operating normally ($s_j = 0$), so that maintenance is not needed ($f_j = 0$). Constraint (6) ensures that the DV to fix BSs is set to 1 iff $s_j = 1$. Constraint (7) guarantees that total number of fixed BSs is less than a pre-defined threshold U (available resources). Constraint (8) ensures that the number of users who can be active at the same time in a BS is no greater than upper bound M_j. Constraint (9) ensures that the transmission power radius of each BS j is between 0 and R_j. Constraint (10) requires that a BS serves its slave MUs under a pre-defined CBP threshold. Constraints (11) and (12) enforce the binary property of the DVs. Finally, Constraint (13) denotes the integer and nonnegative property of DV c_j for channel allocation.

4. Solution approach

4.1. Lagrangean relaxation

The procedure of the Lagrangean relaxation (LR) method is as follows: relax complicating constraints, multiply the relaxed constraints by the corresponding Lagrangean multipliers, and add the relaxed constraints to the primal objective function. The primal optimization problem (IP) can be transformed into the problem (LR) and then into a Lagrangean dual problem. The better primal feasible solution is an upper bound (UB) of the primal problem, while the Lagrangean dual problem solution guarantees the lower bound (LB) of the primal problem. Generally, selecting a level of relaxation is a tradeoff between two properties: the tightness of the gaps between the bounds produced, and the computation time needed to obtain these bounds. To reduce the computational complexity, the constraints that contain at least two decision variables are relaxed, i.e. Constraints (1)–(3), because those DVs can be separated and decided independently after relaxing the respective constraints. In addition, some constraints related to the BS recovery decision, i.e., Constraints (5) and (6), are relaxed. The problem (LR) is further decomposed into three independent subproblems: (SUB 1) channel assignment, power control, and capacity management; (SUB 2) admission control; and (SUB 3) BS recovery. Each subproblem can be optimally and independently solved.

$$
\begin{aligned}
Z_D(v_j^1, v_j^2, v_{jt}^3, v_{jt}^4, v_j^5) = \min \sum_{j \in B''} g_j BB_j(g_j, c_j) + \sum_{t \in T} A_t \left(1 - \sum_{j \in B''} z_{jt}\right) + \sum_{j \in B''} v_j^1 \left(\left(\frac{E_b}{N_{total}}\right)_{req} \right. \\
\left. + \left(\frac{E_b}{N_{total}}\right)_{req} \frac{1}{G} \alpha \frac{P}{N_0} \left((c_j - 1) + \sum_{\substack{j' \in B'' \\ j' \neq j}} \left(\frac{\frac{r_{j'}}{2}}{\text{Max}(D_{jj'} - \frac{r_{j'}}{2}, \varpi)}\right)^\tau c_{j'}\right) - \frac{P}{N_0}\right) + \sum_{j \in B''} v_j^2 \left(\sum_{t \in T} A_t z_{jt} - g_i\right) \\
+ \sum_{t \in T} \sum_{j \in B'} v_{jt}^3(D_{jt} z_{jt} - r_j u_{jt}) + \sum_{t \in T} \sum_{j \in B} v_{jt}^4(u_{jt} - (1 - f_j)(1 - s_j) - f_j \cdot s_j) + \sum_{j \in B} v_j^5(f_j - s_j) \quad \text{(LR)}
\end{aligned}
$$

subject to: (4), (7)–(13).

Subproblem (SUB 1) related to DVs c_j, r_j, and g_j:

$$
\begin{aligned}
Z_{SUB1} = \min \sum_{j \in B''} \left(g_j BB_j(g_j, c_j) + \left(\frac{E_b}{N_{total}}\right)_{req} \frac{1}{G} \alpha \frac{P}{N_0} v_j^1 \left(c_j + \sum_{\substack{j' \in B'' \\ j' \neq j}} \left(\frac{\frac{r_{j'}}{2}}{\text{Max}(D_{jj'} - \frac{r_{j'}}{2}, \varpi)}\right)^\tau c_{j'}\right) \right. \\
\left. + \sum_{t \in T} (v_{jt}^4 - v_{jt}^3 r_j) u_{jt} - v_j^2 g_j + v_j^1 \left(\left(\frac{E_b}{N_{total}}\right)_{req} - \frac{P}{N_0} - \left(\frac{E_b}{N_{total}}\right)_{req} \frac{1}{G} \alpha \frac{P}{N_0}\right)\right) \quad \text{(SUB 1)}
\end{aligned}
$$

subject to: (8)–(10), and (13).

Problem (SUB 1) can be decomposed into $|B''|$ subproblems, since the values of DVs r_j and c_j are discrete and limited. To get an optimal solution, we exhaustively search for all possible combinations of c_j, r_j and g_j.

Subproblem (SUB 2) related to DV z_{jt}:

$$Z_{\text{SUB2}} = \min \sum_{t \in T} A_t \left(1 - \sum_{j \in B''} z_{jt}\right) + \sum_{j \in B''} v_j^2 \sum_{t \in T} A_t z_{jt} + \sum_{t \in T} \sum_{j \in B''} v_{jt}^3 D_{jt} z_{jt}$$

$$= \min \sum_{t \in T} \sum_{j \in B''} (v_{jt}^3 D_{jt} + (v_j^2 - 1)A_t) z_{jt} + \sum_{t \in T} A_t \tag{SUB 2}$$

subject to: (4) and (11).

In (SUB 2), the second term $\sum_{t \in T} A_t = |T| \times A_t$ is a constant value, i.e., the total aggregate traffic in the system. This can be dropped initially and added to the optimal value later, since it will not affect the optimal solution of (SUB 2). The subproblem can then be decomposed into $|B''| \times |T|$ independent subproblems for each $l_{jt} = (v_{jt}^3 D_{jt} + (v_j^2 - 1)A_t)$ where $j \in B''$ and $t \in T$. To derive the minimal value of (SUB 2), we assign either $z_{jt} = 1$ to all $l_{jt} \leqslant 0$, or $z_{jt} = 1$ as the minimal value of l_{jt} for all $|B''| \times |T|$ subproblems.

Subproblem (SUB 3) related to f_j:

$$Z_{\text{SUB3}} = \min \sum_{t \in T} \sum_{j \in B} v_{jt}^4(-(1 - f_j)(1 - s_j) - f_j \cdot s_j) + \sum_{j \in B''} v_j^5(f_j - s_j)$$

$$= \sum_{j \in B} \left(\sum_{t \in T} v_{jt}^4(1 - 2s_j) + v_j^5\right) f_j + \sum_{j \in B} \left(\left(\sum_{t \in T} v_{jt}^4 - v_j^5\right) s_j - \sum_{t \in T} v_{jt}^4\right) \tag{SUB 3}$$

subject to: (7) and (12).

In (SUB 3), the second term $\sum_{j \in B}\left(\left(\sum_{t \in T} v_{jt}^4 - v_j^5\right)s_j - \sum_{t \in T} v_{jt}^4\right)$ is a constant, which can be calculated easily. Let $k_j = \sum_{t \in T} v_{jt}^4(1 - 2s_j) + v_j^5$ so that (SUB 3) can be decomposed into $|B|$ independent subproblems for each k_j. To obtain the minimal value of (SUB 3), we assign at most $f_j = 1$ for the first $|U|$ minimal value of k_j (Constraint (7)), and $f_j = 0$ otherwise.

4.2. Lagrangean dual problem and subgradient method

According to the weak Lagrangean duality theorem [11], for any $(v_j^1, v_{jt}^3, v_{jt}^4, v_j^5) \geqslant 0$ and v_j^2, the objective value of $Z_D(v_j^1, v_j^2, v_{jt}^3, v_{jt}^4, v_j^5)$ is a lower bound on Z_{IP}. Based on the problem (LR), the following dual problem (D) is constructed to calculate the tightest lower bound (LB).

$$Z_D = \max Z_D(v_j^1, v_j^2, v_{jt}^3, v_{jt}^4, v_j^5) \tag{D}$$

subject to: $(v_j^1, v_{jt}^3, v_{jt}^4, v_j^5) \geqslant 0$ and v_j^2.

Then, a subgradient method is applied to solve the dual problem. Let the vector S be a subgradient of $Z_D(v_j^1, v_j^2, v_{jt}^3, v_{jt}^4, v_j^5)$ at $(v_j^1, v_j^2, v_{jt}^3, v_{jt}^4, v_j^5)$. In iteration k of the subgradient optimization procedure, the multiplier vector π is updated by $\pi^{k+1} = \pi^k + t^k S^k$, in which t^k is a step size determined by $t^k = \delta(Z_{IP}^* - Z_D(\pi^k))/\|S^k\|^2$, where Z_{IP}^* is an UB on the primal objective function value after iteration k; and δ is a constant, where $0 \leqslant \delta \leqslant 2$.

4.3. Getting primal feasible solutions

Based on the problem (LR), the dual problem $Z_D = \max Z_D(v_j^1, v_j^2, v_{jt}^3, v_{jt}^4, v_j^5)$ is constructed to calculate the tightest LB subject to $(v_j^1, v_{jt}^3, v_{jt}^4, v_j^5) \geqslant 0$ and v_j^2. We have developed Algorithm A to get primal feasible solutions.

Algorithm A

Step 1. Check the decision feasibility of BS recovery by Constraint (6). Reset $f_j = 0$ if it was $f_j = 1$ in problem Z_D for a workable state of BS j; otherwise, go to Step 2.

Step 2. Recalculate indication function μ_{jt} if the process of DV f_j was adjusted in Step 1.

Step 3. Check QoS Constraint (1) for each BS j.
    Step 3.1. Reduce the transmission power radius r_j if the QoS constraint is still violated; otherwise go to Step 3.2.
    Step 3.2. Compute the aggregate traffic flow g_j based on the transmission power radius r_j determined in Step 3.1.

Step 4. Check the CBP Constraint (10) $BB_j(g_j, c_j) \leqslant \beta_j$. If it is still violated, increase the available channels c_j to meet the requirement β_j; otherwise, go to Step 5.

Step 5. Adjust the transmission power radius r_j so that it just covers all MUs in each BS j.

Step 6. Calculate the total blocked traffic in the system based on the DVs, including c_j, r_j, g_j, and z_{jt}, solved in the previous steps.

Step 7. End algorithm.

5. Computational experiments

5.1. Experiment environment

The parameters used in the experiment are as follows [12,13]: $P/N_0 = 7$ db, $E_b/N_{\text{total}} = 6$ db, $M_j = 120$, $G = 156.25$, $A_t = 0.11$, $\alpha = 0.75$.

The algorithm implementation is coded in C programming language and the running platform is a PC with an INTEL P4-1.6 GHZ CPU. We evaluate the algorithm for 5 cases of BS ($|B|$): 40, 80, 120, 160, and 200; and 6 cases of MU ($|T|$): 500, 1000, 1500, 2000, 2500, and 3000. The number of failed BSs ($|F|$) accounts for one-tenth of all BSs, and the BS recovery ratios (BSRR) are 0.00, 0.25, 0.50, and 0.75. The limit of iterations for the Lagrangean relaxation approach is 1000, and the improvement counter is 25. The parameter δ adopted in the subgradient method is initialized to 2 and halved when the dual objective function value does not improve for 25 iterations. In addition, a pre-defined CBP threshold for each BS is given as $\beta_j = 0.03$.

5.2. Performance analysis

The analysis of the scalability of the extended model is based on $|T| = 3000$ (Fig. 3) and $|B| = 200$ (Fig. 4). A number of $|B|$ and $|T|$ are compared in these figures. Our experiments calculate all near-optimal solutions with gaps of less than 0.001%, and the results include the CBR (UB of Z_{IP}), the service rate, and the CPU time. We now discuss the results in detail.

(1) CBR (R_B): In the case of $|T| = 3000$, the difference in R_B is insignificant for a number of $|B|$, as shown in Fig. 3(a). In the case of $|B| = 200$ in Fig. 4(a), R_B is a monotonically increasing function of $|T|$, which affects R_B significantly. In Fig. 3(b), the case of $|T| = 3000$, the increase in R_B varies between 15% and 71% in the case of $|B| = 40$, while it is in the range 10–44% in the case of $|B| = 200$. *The BSRR is a key factor in reducing R_B for a small number of BSs*. In Fig. 4(b), given $|B| = 200$, the maximum increase of R_B is 27% in the case of $|T| = 500$, while it is 44% in the case of $|T| = 3000$. *The BSRR is an important factor in reducing R_B when the system is heavily loaded*.

(2) Service rate (R_S): Obviously, R_S is a monotonically increasing function of the BSRR. However, the BSRR has an insignificant impact on R_S in all cases of $|B|$ and $|T|$ in Figs. 3(c) and 4(c), respectively. Irrespective of which case is calculated, R_S is in the range 0.92–0.98.

(3) CPU time: Even though our experiments run the LR method up to 1000 iterations, the experiments converge in less than 1000 iterations. Since survivability analysis focuses on available BSs, the more
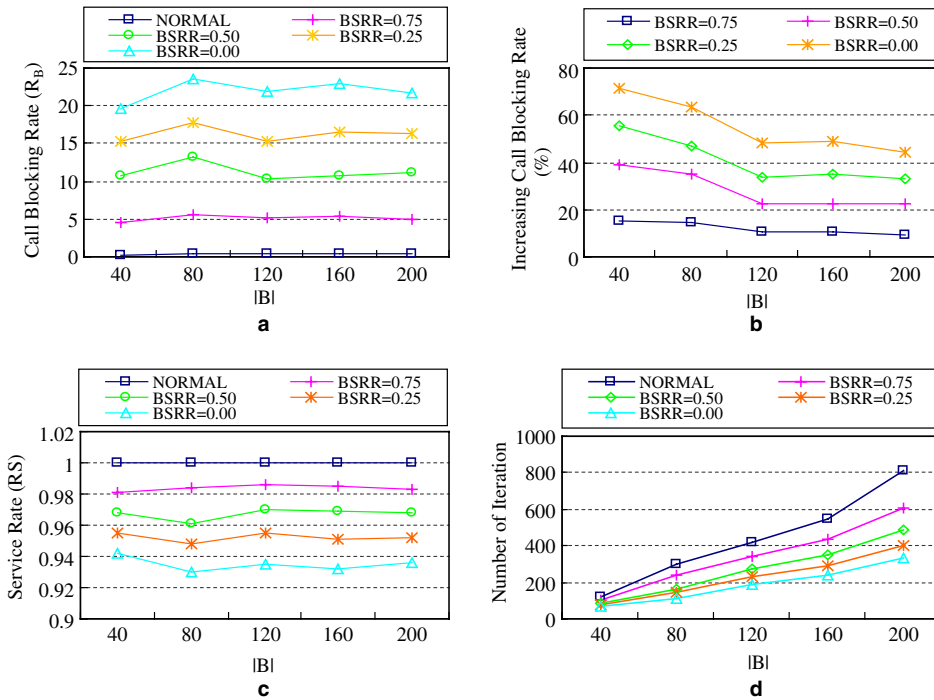
Fig. 3. Analysis of the extended survivability model, with $|T| = 3000$: (a) CBR with respect to BSRR; (b) increasing CBR compared with the normal case; (c) CBR with respect to BSRR; (d) CPU time with respect to BSRR.
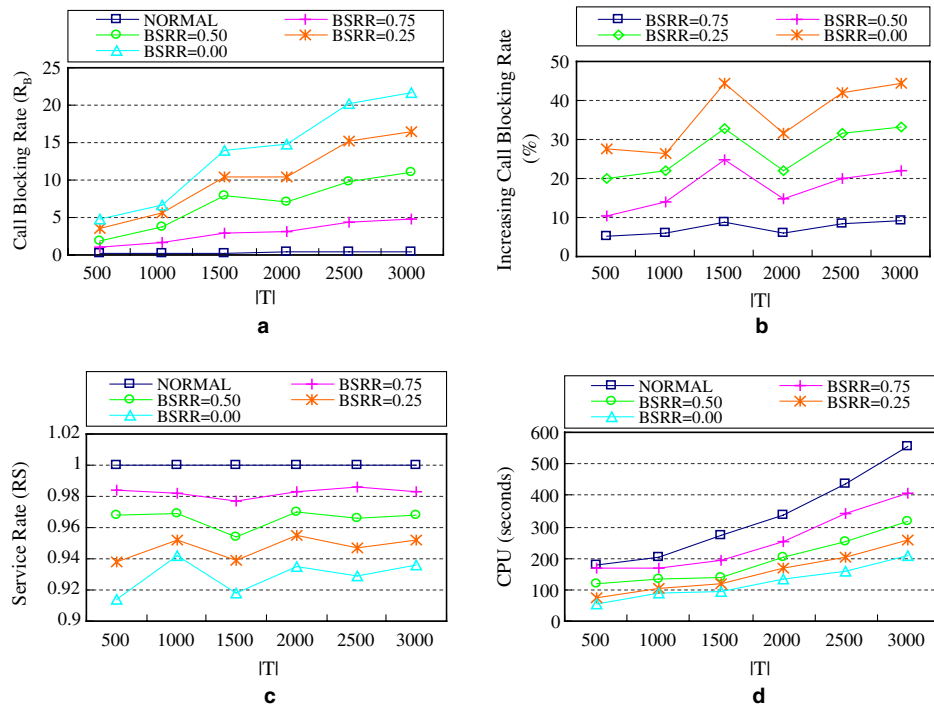

Fig. 4. Analysis of the extended survivability model, given $|B| = 200$: (a) CBR with respect to BSRR; (b) increasing CBR compared with normal case; (c) CBR with respect to BSRR; (d) CPU time with respect to BSRR.

Table 5
Comparison of survivability models in terms of scalability

Model	$|T|/|B|/|F|$	CPU/$C(|F|, |F'|)$		
		BSRR = 0.25	BSRR = 0.5	BSRR = 0.75
Basic	1000/16/4	344/4	517/6	346/4
Extended	3000/200/4	250/20	310/190	400/1140

BSRR required, the more CPU time consumed. In the case of $|T| = 3000$ in Fig. 3(d), the CPU time consumed is between 8 and 16 s in the smallest case $|B| = 40$, while it is between 209 and 556 s in the largest case $|B| = 200$. In the case of $|B| = 200$ in Fig. 4(d), the CPU time consumed is between 57 and 117 s in the smallest case $|T| = 500$, while it is between 209 and 556 s in the largest case $|T| = 3000$. The time consumed is more significant when the number of BSs increases, than when the number of MUs increases. Comparison of the two models demonstrates the flexibility of the modeling and that the extended model is capable of solving the scalability problem. Table 5 compares the CPU time of the two models, where the problem size is given by $|T|$, $|B|$, and $|F|$; and $|F'|$ is defined by the BSRR, i.e., $|F'| = |F| \times \text{BSRR}$. The total number of candidate BSs to be fixed is calculated by $C(|F|, |F'|)$.
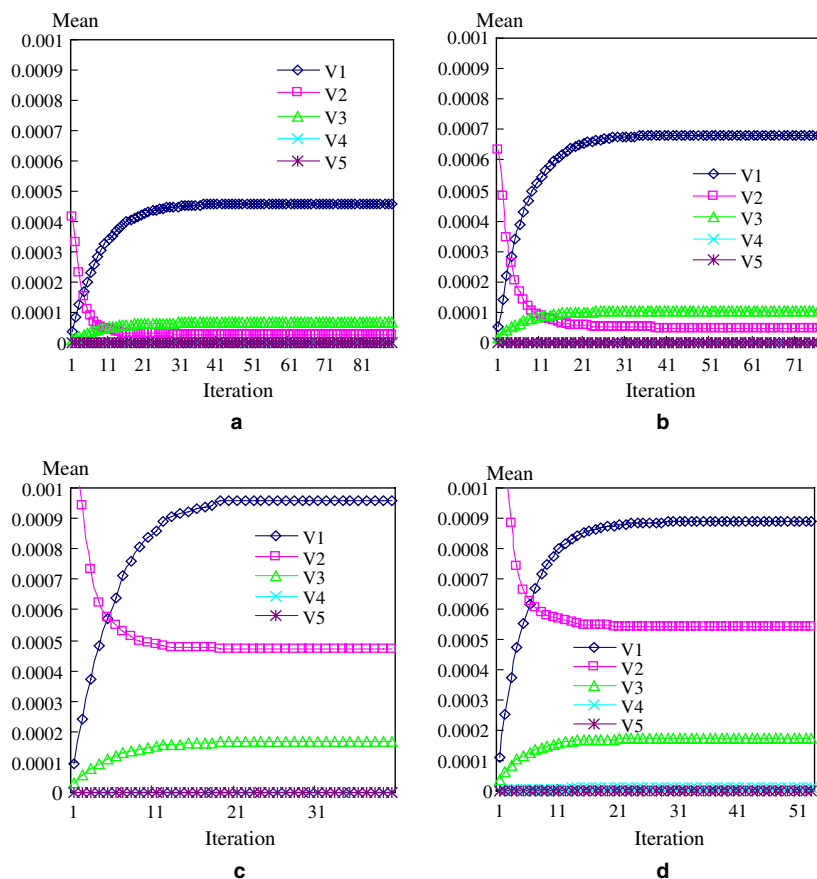

Fig. 5. Sensitivity analysis of the Lagrangean multipliers for network survivability with different BSRRs: (a) BSRR = 0.75, (b) BSRR = 0.5, (c) BSRR = 0.25, (d) BSRR = 0.
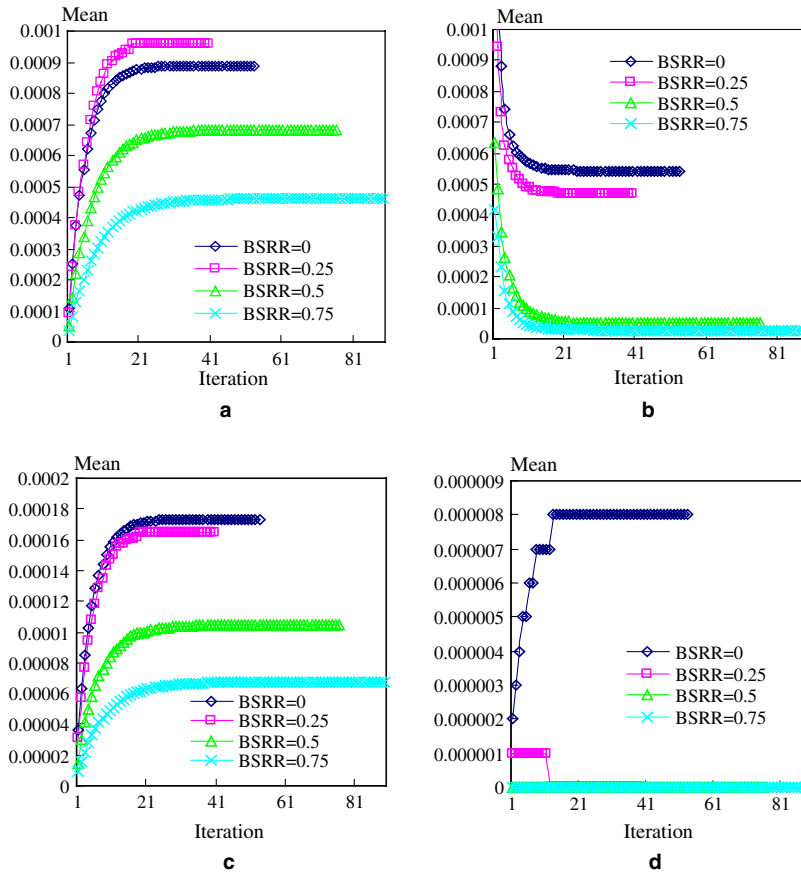
Fig. 6. Sensitivity analysis of BSRR for network survivability given a specific multiplier: (a) V1, (b) V2, (c) V3, (d) V4.

6. Sensitivity analysis

In this section, given $|B| = 36$, $|T| = 1000$, we analyze the mean value of the following multipliers: V1 related to Constraint (1), V2 related to Constraint (2), V3 related to Constraint (3), V4 related to Constraint (5), and V5 related to Constraint (6). All values of V5 are calculated close to zero. Fig. 5 shows the effect of the BSRRs on the values of the multipliers, which are functions of the number of iterations. The calculations converge with different numbers of iterations. Nevertheless, the multipliers are all increasing functions of the number of iterations, except for multiplier V2, because it can be a negative value. The results show that the QoS constraint (V1) is the most important in the survivability problem, especially when the BSRR is 0.25. This finding is validated by Fig. 6. In Fig. 6(a), the V1 value converges to 0.01 (the largest value) with respect to BSRR = 0.25 (vs. other BSRRs). The CBR is affected by the constraint of checking the aggregate flow (V2), as in Fig. 6(b). This constraint becomes much more important with a smaller BSRR (0–0.25) than a larger BSRR (0.5–0.75). The coverage constraint (V3) plays a significant role when no BS is recovered, as shown in Fig. 6(c).

We also investigate the effects of the problem of scalability. Given $|B| = 200$ in Fig. 7, irrespective of which BSRR is considered, the QoS constraint (V1) is critical when the system has a heavy load ($|T|$ larger than 2500) than when there is a light load ($|T|$ less than 2500). The more $|T|$ there are in the system, the larger the value obtained. When the system's load is light, the constraint for calculating the aggregate flow is more important than the QoS constraint. This means that it is more important to minimize the CBR for a light load than for a heavy load.

In Fig. 8, if $|T| = 3000$ is given, the QoS constraint (V1) is a significant factor. Nevertheless, the multiplier value decreases substantially when $|B|$ increases, because the average system load decreases. The sensitivity analysis yields the same results as the experimental analysis.
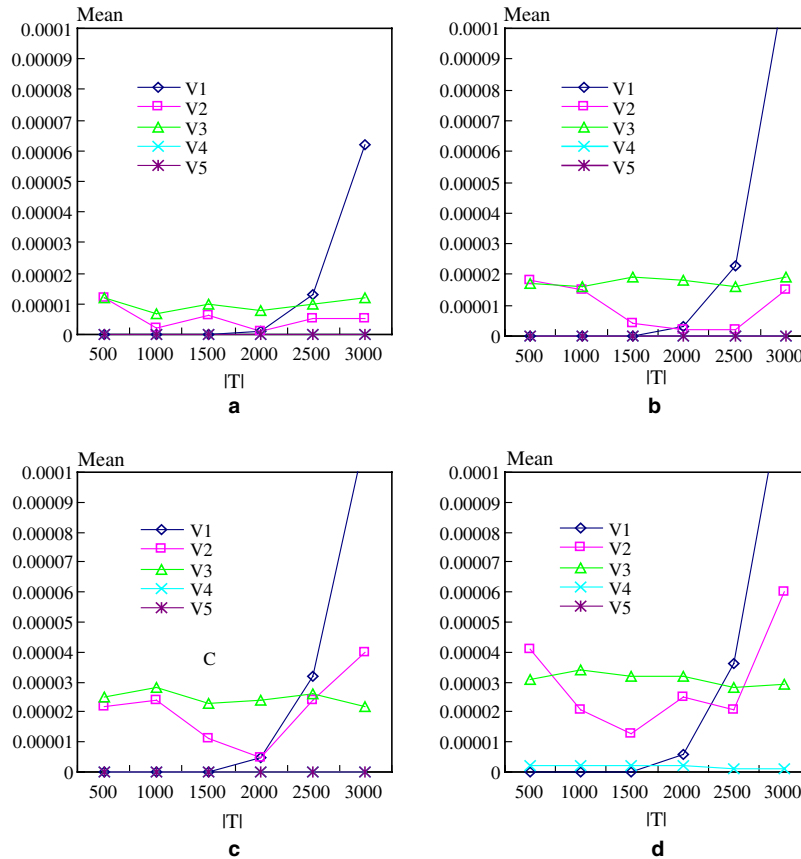
Fig. 7. Sensitivity analysis of network survivability given $|B| = 200$ and different BSRRs. The Lagrangean multiplier is a function of $|T|$: (a) BSRR = 0.75, (b) BSRR = 0.5, (c) BSRR = 0.25, (d) BSRR = 0.

7. Concluding remarks

7.1. Research contributions

Survivability is always a crucial issue in both wireline and wireless networks. In the case of BS failures, when maintenance is undertaken to ensure a network's performance, we propose a flexible survivability model to relocate spare resources in such a way that total call blocking is minimized. In other words, we consider the survivability issue at the operational level. We reformulate the basic model to reduce the time complexity from $O(N^2)$ to $O(1)$, where N is the number of failed BSs. This indicates that the process of mathematical modeling is an issue of concern in problem solving. The goal of the extended survivability model is to offer survivable services that minimize total system call blocking. In addition, this model can also be applied to dynamic capacity expansibility. Comparison of the basic and extended models proves the flexibility of our modeling, and demonstrates that the extended model can deal with the scalability problem by allocating portable/mobile BSs anywhere when needed.

7.2. Solution approach advantages

We apply LR as our solution approach because the complicated optimization problem discussed in this paper is NP-hard. Even though LR is a standard solution technique that can solve a wide range of combinatorial optimization problems, to efficiently and effectively solve those problems, the following nontrivial and novel tasks/stages should also be considered: (1) the problem formulation stage: a suitable formulation that
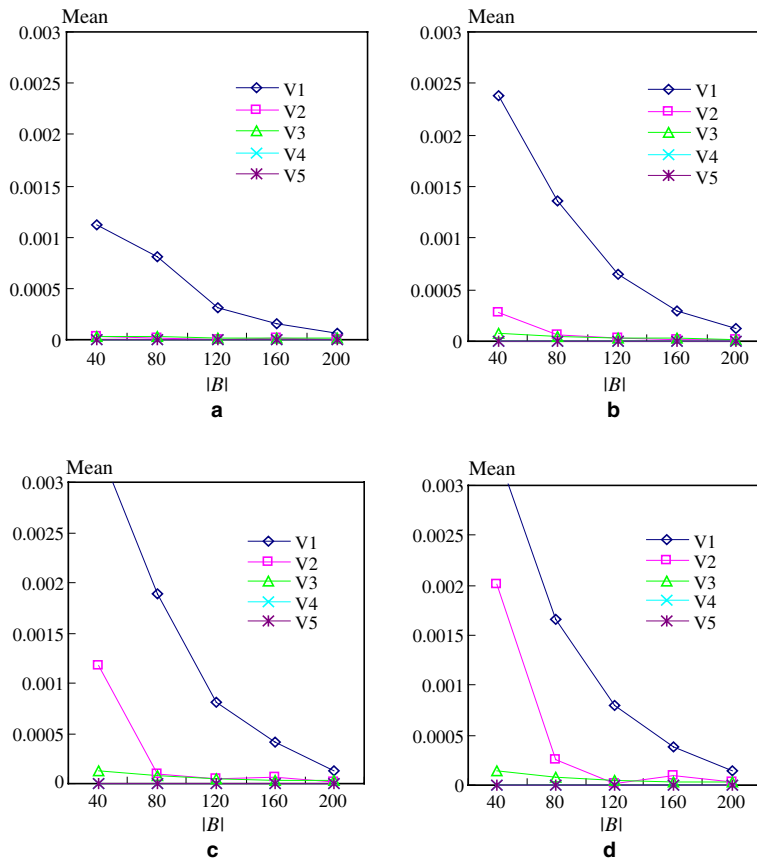
Fig. 8. Sensitivity analysis of network survivability given $|T| = 3000$ and different BSRRs. The Lagrangean multiplier is a function of $|B|$: (a) BSRR = 0.75, (b) BSRR = 0.5, (c) BSRR = 0.25, (d) BSRR = 0.

can be decomposed into a number of subproblems to which LR can be successfully applied. This may require many attempts to reformulate the problem by trial and error; (2) the solution procedure stage: decides which constraints should be relaxed and how a number of the critical parameters can be carefully determined by a procedure, so that the optimal solution and a high convergence rate can be achieved; (3) The primal feasible solution: how to apply Lagrangean multipliers and develop an efficient algorithm to get primal feasible solutions is a challenging issue. We use Lagrangean multipliers for sensitivity analysis so that the corresponding constraints can be evaluated for decision support.

7.3. Managerial implications

The total system CBR is much less affected by the CBP threshold of each BS when the network load is light than when it is heavy; hence, the more traffic loaded, the less the service rate varies. BS recovery is much more important in a light load than in a heavy load. The BSRR is a key factor in reducing the blocking rate for a small number of deployed BSs, and it is the most important factor for reducing the blocking rate when the network load is heavy. The proposed extended model provides a survival service, subject to the allocation of available resources.

References

[1] Medhi D. A unified approach to network survivability for teletraffic networks: models, algorithms and analysis. IEEE Trans Commun 1994;42(2–4):534–48.

[2] Medhi D, Khurana R. Optimization and performance of network restoration schemes for wide-area teletraffic networks. J Network Syst Manage 1995;3(September):265–94.

[3] Clarke LW, Anandalingam G. A bootstrap heuristic for designing minimum cost survivable networks. Comput Oper Res 1995;22(9):921–34.

[4] Koh SJ, Lee CY. A tabu search for the survivable fiber optic communication network design. Comput Indust Eng 1995;28(4):689–700.

[5] Ghashghai E, Rardin RL. Using a hybrid of exact and genetic algorithms to design survivable networks. Comput Oper Res 2002;29(1):53–66.

[6] Snow AP, Varshney U, Malloy AD. Reliability and survivability of wireless and mobile networks. IEEE Comput 2000;33(7):49–55.

[7] Chuprun S, Bergstrom CS. Comparison of FH/CDMA and DS/CDMA for wireless survivable networks. In: Proc IEEE GLOBECOM, vol. 3, 1998. p. 1823–7.

[8] Chen AC. Overview of code division multiple access technology for wireless communications. In: Proc IEEE 27th IECON, Aachen, Germany, 1998. p. T15–24.

[9] Chu K-C, Lin FY-S, Lee S-H. Integrated planning and capacity management of survivable DS-CDMA networks. In: Proc IEEE ICNSC, vol. 2, 2004. p. 1154–9.

[10] Chu K-C. Network survivability and performance modeling in cellular communication systems. GESTS Int Trans Comput Sci Eng 2005;8(1):13–24.

[11] Fisher ML. The Lagrangian relaxation method for solving integer programming problems. Manage Sci 1981;27:1–18.

[12] Hernandez MA, Janssen GJM, Prasad R. Uplink performance enhancement for WCDMA systems through adaptive antenna and multiuser detection. In: Proc IEEE VTC-Spring, vol. 1, 2000. p. 571–5.

[13] Tam W-M, Lau FCM. Analysis of power control and its imperfections in CDMA cellular systems. IEEE Trans Vehic Technol 1999;48(5):1706–17.

Kuo-Chung Chu received his B.S. and M.S. degrees in Computer Science from Feng-Chia University, Taiwan, in 1988 and 1990, respectively; and his Ph.D. degree in Information Management from the Department of Information Management, National Taiwan University in 2005. After graduating from the FCU, he joined the Computer Centre, Academia Sinica, Taiwan, where he was responsible for network systems management. In 1994, Professor Chu joined the Faculty of Information Management, Jin-Wen Institute of Technology, Taipei as a lecturer. Since 2005, he has been an Associate Professor in that department. His research interests include decision modeling, simulation, network planning and optimization, management, and performance evaluation of mobile wireless communications networks.

Professor Frank Yeong-Sung Lin received his B.S. degree in Electrical Engineering from the Department of Electrical Engineering, National Taiwan University in 1983; and his Ph.D. degree in Electrical Engineering from the Electrical Engineering Department, University of Southern California in 1991. After graduating from the USC, he joined Telcordia Technologies (formerly Bell Communications Research, abbreviated as Bellcore) in New Jersey, USA, where he was responsible for developing network planning and capacity management algorithms for a wide range of advanced networks. In 1994, Prof. Lin joined the Faculty of Electronic Engineering, National Taiwan University of Science and Technology. Since 1996, he has been with the Department of Information Management, National Taiwan University. His research interests include network optimization, network planning, performance evaluation, high-speed networks, wireless communications systems, distributed algorithms, and information security.