Performability Analysis of Channel Allocation with Channel Recovery Strategy in Cellular Networks

Yue Ma, Student Member, IEEE, Cheul Woo Ro, and Kishor S. Trivedi, Fellow, IEEE

Abstract: In this paper, we propose and compare three channel recovery schemes for fixed channel assignment. In Scheme I, a failed channel is switched by an idle channel whenever it is available. In Scheme II, the switching strategy is employed only after an attempt to restore the failed channel within a short period of time is unsuccessful. Scheme III is a hybrid of Schemes I and II: for some of the failed channels, Scheme I is applied; for the others, Scheme II is used. To reflect the traffic pattern in a realistic way, a fixed point iteration scheme is applied to capture the dynamic behavior of the handoff arrivals. Stochastic reward nets (SRNs) are developed for the performability analysis of the three schemes.

1 Introduction

With the rapid growth in the demand for personal communications services (PCS), efficient channel allocation plays a vital role in providing the pre-specified Quality of Service (QoS). Many channel allocation schemes have been proposed by several researchers [6, 9, 11, 15]. These schemes generally assume a perfect situation where the channel in use never fails. In a practical environment, cellular networks, like any other physical system, are subject to failures. A channel may fail to operate properly due to various reasons such as power loss, software/hardware problems or hostile action. With the increasing demand for cellular communications, a disruption in service could cause severe consequences in both economic and social sense. Thus providing restoration subsequent to channel failures has become an important issue in ensuring network integrity.

To obtain realistic performance measures for cellular networks, one should consider changes in performance due to failure related behavior. In performability analysis [8], simultaneous consideration is given to both performance and reliability/availability measures. The purpose of this paper is to analyze the performability of cellular networks under different channel recovery schemes.

In cellular networks, a given geographical area is divided into a certain number of cells. For the fixed channel assignment (FCA), each cell is allocated a fixed set of duplex channels which are different from those assigned to the adjacent cells so that the mutual interference is beneath the tolerable threshold. When a new call (NC) is attempted in a cell covered by a base station (BS), the NC is connected if a channel is available in the cell. Otherwise, the call is blocked. When a mobile station (MS) travels across the cell boundaries, the channel in the old serving cell is released, and an idle channel is required in the target cell, which would be the new serving cell. This process is called handoff. If an idle channel exists in the target cell, the handoff call (HC) continues nearly transparently to the user. Otherwise, the HC is blocked.

In this paper, we propose and compare three channel recovery schemes. We assume that the system uses the FCA scheme and a fixed number of guard channels are reserved for HCs. In Scheme I, a failed channel (FCh) is automatically switched by an idle channel if one is available. Otherwise, the failed call (FC) is queued until an idle channel is available. In Scheme II, in order to use the limited spectrum more efficiently, a restoration of the FCh is first attempted within a pre-specified time threshold. If the FCh can be recovered within the time threshold, the FC is restored without costing additional channel resources. Otherwise, the switching strategy as in Scheme I is applied. In Scheme III, a hybrid of Schemes I and II is applied: for some of the failed channels, Scheme I is applied; for the others, Scheme II is used.

With different NC arrival rates, the corresponding HC arrival rates vary accordingly. To capture this dynamic behavior, a fixed point iteration scheme [7] is applied to determine the HC arrival rates. Stochastic reward net (SRN) models are developed for the performability analysis. For all the three schemes, the expected blocking probabilities for both NCs and HCs are obtained. To compute the system availability, the expected numbers of available channels are also found. To compare the restoration speeds, the expected recovery times for the FCS are obtained as well.

This paper is organized as follows. In Section 2, we give a brief introduction to SRNs. In Section 3, system and traffic specifications are made for a generic single cell in a cellular network. In Section 4, we describe the channel recovery schemes. In Section 5, we first discuss several measures of performability and give a brief introduction to the fixed point iteration scheme, then the numerical
results for the performance measures are obtained. Finally, we make our conclusions in Section 6.

2 Introduction to SRN

Stochastic reward net [2] is an extension of Petri net (PN) [10], which is a high level description language for formally specifying complex systems. A PN is a bipartite directed graph with two types of nodes: places and transitions. Each place may contain an arbitrary (natural) number of tokens. For a graphical presentation, places are depicted as circles, transitions are represented by bars and tokens are represented by dots or integers in the places. Each transition may have zero or more input arcs, coming from its input places; and zero or more output arcs, going to its output places. A transition is enabled if all of its input places have at least as many tokens as required by the multiplicities of the corresponding input arcs. When enabled, a transition can fire and will remove from each input place and add to each output place the number of tokens corresponding to the multiplicities of the input/output arcs. A marking depicts the state of a PN which is characterized by the assignment of tokens in all the places.

Generalized stochastic Petri nets (GSPNs) [1] extend the PNs by assigning a firing time to each transition. Transitions with exponentially distributed firing times are called timed transitions while the transitions with zero firing times are called immediate transitions. A marking in a GSPN is called vanishing if at least one immediate transition is enabled; otherwise it is called a tangible marking. Under the condition that only a finite number of transitions can fire in finite time with non-zero probability, it can be shown that a given GSPN can be reduced to a homogeneous continuous time Markov chain (CTMC) [1]. GSPN also introduces inhibitor arcs. An inhibitor arc from a place to a transition disables the transition if the place contains at least as many tokens as the cardinality of the inhibitor arc. Graphically, an inhibitor arc is represented by a line terminated with a small circle.

In order to make more compact models of complex systems, several extensions are made to GSPN, leading to the SRN. One of the most important features of SRN is its ability to allow extensive marking dependency. In an SRN, each tangible marking can be assigned one or more reward rates. Parameters such as the firing rate of the timed transitions, the multiplicities of input/output arcs and the reward rate in a marking can be specified as functions of the number of tokens in any place in the SRN. For an SRN, all the output measures are expressed in terms of the expected values of the reward rate functions. To get the performance and reliability/availability measures of a system, appropriate reward rates are assigned to its SRN. In this paper, we use the tool SPNP [3] to specify and solve the SRN models.

3 System Description

Under the condition that all the neighboring cells are statistically identical and behave independently, the characteristics of the overall system can be captured by focusing on a single cell. In our system, all cells are treated equivalently so that we can concentrate on the performance aspects of the channel recovery schemes. By varying the traffic parameters, the presented schemes can also be applied to the cases where cells are located at the border of the covered area.

Suppose that a set of C duplex channels are assigned to a cell. At each BS, C_h out of C channels are reserved exclusively for the HCs. Thus, an arriving NC will be blocked if C - C_h channels are already occupied in a cell. We assume that the arrival processes of NCs and HCs are Poisson with rates λ_n and λ_h respectively. Thus, the total call arrival rate in a given cell is λ = λ_n + λ_h. The channel holding time, which is terminated either by the call clearing or by the MS leaving for another cell, can be described by two exponential distributions with rates λ_d and λ_h respectively [5]. Because of the memoryless property of the exponential distribution [12], the mean call duration (1/λ_d) remains the same for an HC as that of an NC.

We also assume that an idle channel is always in perfect condition for service. In other words, a channel can only fail when it is in service. A call fails when the channel it holds fails. An FC can be restored when an idle channel is switched to it or the FCh is recovered within the pre-specified threshold.

4 Model Description

Scheme I: This scheme is quite similar to the Automatic Protection Switching (APS) scheme which is widely used to enhance the network integrity in the ATM networks [14]. In APS systems, a failed network component is switched by an identical spare component when the protection switch detects a failure. Since the spectrum is limited, no spare channels are reserved for the FCs. However, the FCs are treated with the same priority as the HCs in the sense that both of them can access any available channel in the BS.

In Figure 1, we show an SRN model for this scheme. Place CP is the channel pool for the cell. Initially, there are C idle channels which are accessible for both the NCs and the HCs. Transitions t_n and t_h represent the arrivals of NCs and HCs respectively. Transition t_h is enabled with at least one idle channel in place CP. Otherwise, it is blocked. Transition t_n is disabled if there are less than C_h + 1 channels in place CP. This is represented by the multiple input arc from place CP to transition t_n and the multiple output arc from transition t_n to place CP. The resulting effect is that when transition t_n is enabled, only one token is moved from place CP to place T. The number of tokens in place T is the number of channels currently being utilized in the cell. Transitions t_d and t_h respectively represent the departure of a call, either due to the termination of the call or due to the MS leaving the cell. The clearing rate for a single call is λ_d. The rate at which an MS leaves the cell is λ_h. Transition t_f represents the failure of a channel while it is in use. The failure rate
for a single channel is $\lambda_f$. Notice that transitions $t_f$, $t_d$ and $t_f^1$ have marking dependent firing rates. The firing rates for transitions $t_f^1$, $t_d$ and $t_f$ are $k\lambda_f^1$, $k\lambda_d$ and $k\lambda_f$ respectively, where $k$ is the number of tokens in place $T$. The marking dependency is indicated by the $#$ signs next to the transitions in Figure 1.

![Figure 1: SRN for Scheme I](image)

When a channel fails, the FCh is switched by an idle channel if one is available. In this case, the FC is restored to service immediately. When the channel pool is empty, the FC is queued in the buffer $B$. As soon as an idle channel is available, an FC is restored instantly. The queued FChs are served by the first-in/first-out (FIFO) policy. The above process is represented by the immediate transition $t_1$ in Figure 1. In order to get transition $t_1$ fired, at least one token is required in both places $CP$ (an idle channel) and $B$ (an FC). In the mean time, the FChs are being recovered in place $R$ under the FIFO policy with a single recovery facility. This is represented by transition $t_r$ and the recovery rate is $\mu_r$.

![Figure 2: SRN for Scheme II](image)

**Scheme II:** The SRN for Scheme II is shown in Figure 2. In order to use the limited spectrum more efficiently, a recovery of an FCh is attempted in two phases. For the first phase, a short time threshold is pre-specified. In this phase, the FC is on hold. If the FCh can be restored within this phase, the FC is recovered to the service place $T$. Neither phase two nor additional idle channel is needed. The branching of the recovered and unrecovered FChs after phase one is represented by the temporary place $T_3$ and the immediate transitions $t_3$ and $t_4$. An FCh can be recovered with probability $c$ after completing phase one. With probability $\bar{c} = 1 - c$, an FCh needs a longer recovery time in phase two. The automatic switching strategy in Scheme I is applied in phase two, where the recovery rate is assumed to be the same as in Scheme I.

The short restoration time in phase one is assumed to be a uniform distribution $U(0, b)$, where $b$ is the prescribed short restoration threshold. SRNs do not currently provide transitions with uniformly distributed firing times. However, a non-Markovian distribution can be Markovized by approximating the distribution with the time to absorption of a Markov chain [4]. A uniform distribution can be approximated by a 3-stage Erlang distribution where the transition rate for each stage is $b/3$ [13].

![Figure 3: SRN for Scheme III](image)

**Scheme III:** In order to combine the advantages of Schemes I and II, a hybrid approach of the two schemes is applied in Scheme III. That is, $c_l$ fraction of the FChs are restored with the approach applied in Scheme I. The remaining $\bar{c}_l = 1 - c_l$ fraction are restored using Scheme II. This is represented by the immediate transitions $t_5$ and $t_6$ in Figure 3. When $c_l$ equals 1, Scheme III is equivalent to Scheme I. When $c_l$ equals 0, Scheme III reduces to be Scheme II.

In both Schemes II and III, we assume that there are two recovery facilities. One recovery facility is dedicated to the short recoveries. The other one is responsible for the longer recoveries as in Scheme I.

5 **Measures of Interest and Numerical Results**

We denote the blocking probability for the HC as $P_{bh}$. To calculate $P_{bh}$, the reward rate assignment is:

73
Here $r_{th}^j$ is the reward rate for state $j$ in the CTMC of $P_{bn}$, and $\#(CP_j)$ represents the number of tokens in place $CP_j$ in marking (state) $j$. Thus a reward rate of 1 is assigned to the states where the channel pool is empty, and a reward rate of 0 is assigned to the other states. Then $P_{bh}$ is calculated by

$$P_{bh} = \sum_{j \in \Omega} r_{th}^j \pi_j,$$

where $\Omega$ is the set of all tangible markings and $\pi_j$ is the steady-state probability of marking $j$. To calculate different measures of the system, we need to have different assignments of the reward rates. To get the blocking probability of a NC, which is denoted by $P_{bn}$, a reward rate of 1 should be assigned to the states where $CP_j$ has less than $C_h + 1$ tokens, that is,

$$r_{bn}^j = \begin{cases} 1 & \text{if } \#(CP_j) \leq C_h, \\ 0 & \text{otherwise.} \end{cases}$$

Assignments (1) and (2) are used to obtain $P_{bh}$ and $P_{bn}$ for all the three schemes.

Figure 4: Blocking probabilities for the three schemes.

We use the expected number of available channels to assess the availability of the system. This is achieved by assigning the reward rate equals to the number of tokens in the channel pool, that is, $r_{th}^j = \#(CP_j)$.

Little's formula [12] is applied to get the expected recovery time for an FC which is eventually restored to service. Let $\Lambda_f$ denote the actual throughput of transition $t_f$. For Scheme I, the expected recovery time (EIT) is:

$$E[RT_f] = \frac{E[B]}{\Lambda_f},$$

where $E[B]$ is the expected number of tokens in place $B$. After simplification, the expected recovery times for Schemes II and III are expressed in Equations (4) and (5):

$$E[RT_{II}] = \frac{E[F] + \frac{1}{2}(E[T_1] + E[T_3])}{\Lambda_f} + E[B],$$

$$E[RT_{III}] = \frac{E[F] + \frac{1}{2}(E[T_1] + E[T_3])}{\Lambda_f} + E[B].$$

Numerical results are obtained to show the performance characteristics of the three channel recovery schemes. For the purpose of discussion, we assume the short restoration threshold is 3 s, and 77% of the FCs can be recovered within the pre-specified threshold. The average recovery time $1/\mu_r$ is assumed to be 45 s. We also assume that the average call duration time $1/\lambda_0$ is 120 s. The failure rate $\lambda_f$ for each channel is 0.06793/min. We assign $c_1 = 0.5$ so that Scheme III combines Schemes I and II in a neutral way. The rate $\lambda_0^m$ at which an MS leaves a cell is 0.33/min. The number of total channels assigned to each BS is $C = 20$ and the number of guard channels reserved for handoff calls is $C_h = 1$.

Figure 5: Expected Number of Available Channels

The arrival rate $\lambda_h$ of HCs should be equal to the actual throughput of transition $t_h^j$, denoted by $\Lambda_h^m$. The value of $\Lambda_h^m$ can be calculated as follows:

$$\Lambda_h^m = \sum_{j \in \Omega} \#(T_j) \lambda_0^m \pi_j (\lambda_h^m) = \sum_{j \in \Omega} \#(T_j) \lambda_0^m \pi_j (\lambda_h^m),$$

where $\#(T_j)$ denotes the number of tokens in place $T$ in marking (state) $j$, and $\pi$ is the steady-state probability vector of the SRN models. Notice that $\pi$ is a function of $\Lambda_h^m$, because the firing rate of transition $t_h^j$ should be equal to $\Lambda_h^m$. In [7], Equation (6) is defined as the fixed point equation and $\Lambda_h^m$ is termed as the iteration variable. According to Theorem 2 in [7], a fixed point will exist if 1) the functions of $\pi$ are weighted sums of state probabilities and the weights are constants; 2) the CTMCs underlying the SRNs are irreducible with more than one state. It is easy to verify that the three SRN models that we have developed satisfy the above two conditions. Therefore, a fixed point exists for Equation (6). For all the three schemes, the iteration process converges fairly rapidly, generally, in less than ten iterations.

In Figure 4, we show the blocking probabilities of NCs and HCs with respect to different NC arrival rates. It is shown that both Schemes II and III provide lower blocking probabilities for both NCs and HCs. By introducing a short restoration time in Schemes II and III, the blocking probabilities can be reduced a lot.

From Figure 5, we can see that both Schemes II and III offer more expected available channels than Scheme I does. Thus, from both the performance and the availability perspectives, Schemes II and III offer a better level of
QoS. In Figure 6, we show the expected recovery times of FCs for the three schemes. Because of the short pre-specified threshold, the recovery times for Schemes II and III are longer than those of Scheme I. Table 1 shows the state space for the SRN models of the three Schemes.

Figure 6: Expected Recovery Times for the Three Schemes
Considering Figures 4, 5 and 6, we can see that there is a tradeoff between the performability measures and the recovery times for the FCs. Schemes II and III offer low blocking probabilities with the price of recovering the FCs within one or two seconds and requiring two recovery facilities, whereas Scheme I yields relatively high blocking probabilities, but can restore the FCs within a time period which is virtually undetectable by the user and requires only one recovery facility.

<table>
<thead>
<tr>
<th>Schemes</th>
<th>No. of Tangible Markings</th>
<th>No. of Vanishing Markings</th>
<th>No. of Nonzero Transitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>440</td>
<td>381</td>
<td>2400</td>
</tr>
<tr>
<td>II</td>
<td>9050</td>
<td>17370</td>
<td>74509</td>
</tr>
<tr>
<td>III</td>
<td>9050</td>
<td>27118</td>
<td>92127</td>
</tr>
</tbody>
</table>

Table 1: State Space for the Three Schemes with $C = 20$ and $C_b = 1$
As the limited spectrum gets more crowded with the growing popularity of cellular communications, the additional hardware cost of Schemes II and III is necessary to guarantee the QoS, especially during the peak hours. Scheme III combines the advantages of Schemes I and II in the sense that it provides less waiting time (vs. Scheme II) and offers lower blocking probabilities (vs. Scheme I). By adjusting parameter $c_1$ in Scheme III, the network administrator can manage to get both low blocking probabilities and tolerable recovery times that satisfy the QoS.

6 Conclusion
In this paper, we have developed SRN models to investigate the performability characteristics for three channel recovery schemes. To reflect the traffic pattern in a realistic way, a fixed point iteration scheme is applied to capture the dynamic behavior of the handoff arrivals.

Schemes II and III offer lower blocking probabilities and high channel availabilities with the cost of detectable recovery times for the user and additional hardware. With the increasing demand for cellular communications, additional costs are necessary to make efficient use of the limited spectrum. By adjusting the parameter $c_1$ in accordance with the traffic pattern, Scheme III can offer both low blocking probabilities and tolerably short recovery times. Thus, Scheme III is a good candidate for the channel recovery scheme in cellular networks.

References