Performance Analysis of Microcellisation for Supporting Two Mobility Classes in Cellular Wireless Networks

K. Maheshwari and Anurag Kumar

Abstract— **We study the call blocking performance obtained by microcellising a macrocell network. Each macrocell is partitioned into microcells, and some of the channels originally allocated to the macrocell are assigned to the microlayer cells according to a reuse pattern. The arriving calls are classified as *fast or slow*; fast calls are always assigned only to macrocell channels, whereas for slow calls a microcell channel is first sought. Slow calls may be allowed to *overflow* to the macrolayer, but may be *repacked* to vacated microcell channels. Calls can *change their mobility* class during a conversation. We develop an approximate analysis for computing the slow and fast call blocking probabilities in such a system. We adopt the technique of analysing an isolated macrocell with the Poisson arrival assumption, and then iterating on the stationary analysis of the isolated macrocell to obtain stationary results for the multicell system. Simple, but accurate approximations are developed for analysing the isolated macrocell and its associated microcells. The analyses based on the approximate isolated cell model are validated against *simulations of a multicell model*.**

I. Introduction

In cellular wireless mobile telephony systems, a decrease in the size of the cells allows more frequency reuse in a given area. With the decrease in size of the cells, however, there is an increase in the number of cell boundaries that a mobile unit crosses. These boundary crossings stimulate handoffs and location tracking operations. Thus the signalling capacity of the signalling processors (in the Base Stations and the Mobile Switching Centres) can limit the call handling capacity of a cellular system as the cell size is decreased. These issues are discussed in [7].

One way of controlling the increase of signalling traffic, while deriving the frequency reuse advantage of smaller cells, is to consider a cellular (*macrocellular*) network, and subdivide the large cells into smaller *microcells* (see [14]). Radio channels are allocated to macrocells and to microcells. Each mobile call is then classified as belonging to one of two mobility classes, *fast* and *slow*. A call that originates at or terminates on a slow mobile (henceforth referred to as a *slow call*) is allocated to a channel in the microcell in which the mobile is currently located, whereas a fast call is allocated to a macrocell. It can be expected that, with appropriate engineering of such a system, more traffic can be handled, with a given number of channels and a required grade of service, while limiting the increase of signalling traffic on the network. See [3], [1], [9], [22] for further discussions of such multitier cellular network architectures.

The main contribution of this paper is to develop an approximate analysis for calculating the probabilities of call blocking in a model of a microcellular network; the analysis is verified by simulations of the multicell model. The scenario that we are concerned with is that there is a macrocellular network, with a given frequency allocation to each cell. Each macrocell is then *microcellised*, and the original frequencies assigned to each cell are partitioned between the microcells and the original macrocell[1]. A call that is handled by a channel in a macrocell is said to be in the *macrolayer* while a call that is handled by a channel in a microcell is said to be in the *microlayer*.

For the purpose of this study, we assume that a speed threshold, used for classifying the mobiles, has been determined. A call is identified as fast or slow by the cellular system. Approaches for carrying out such classification are proposed in [10], [13], [22]; we assume, as in [9], that such classification has already been done on call arrival. A fast call is allocated to a macrolayer channel in the macrocell that it is located, and a call that is identified as slow is allocated to a microlayer channel in the microcell that it is located. A call is blocked in a layer if all the channels in that layer are occupied. A slow call that is blocked in the microlayer is attempted to be assigned a channel in the macrolayer. These calls are said to *overflow* from the microlayer to the macrolayer. A slow call is thus blocked in the system only if channels in both the macrocell and the microcell (in which it is located) are occupied. A fast call is blocked if all the channels in the macrocell to which it belongs are occupied. Overflow of slow calls to the macrolayer may give them undue advantage over the fast calls; to reduce this advantage, one possibility is that if there are slow calls in the macrolayer from a particular microcell, one of these calls is moved to the microlayer whenever a call departs from that microcell. We call this procedure *repacking*.

In reality, mobiles do not move with constant speeds. A speed change occurs when a mobile moves from a more crowded area to a less crowded area or if a mobile encounters a traffic signal. This aspect is also included in our model by allowing calls to undergo *mobility change*; i.e., a

K. Maheshwari is currently at Lucent Technologies Bell Laboratories, Holmdel, N.J., 07733, USA; email: mahe@lucent.com

Anurag Kumar is with the Dept. of Electrical Communication Engg., Indian Institute of Science, Bangalore, 560012, INDIA; e-mail: anurag@ece.iisc.ernet.in. A part of his work on this paper was done during his sabbatical at the Wireless Information Network Laboratory (WINLAB), Rutgers University, N.J.

[1]Clearly, there are, other, more efficient, channel allocation schemes, and our analysis approach applies to any static allocation scheme. The particular allocation that we have described here is perhaps the first that a cellular operator may adopt, as it does not disturb an already established frequency plan.

fast call can become a slow call and vice-versa.

We make the standard stochastic assumptions; i.e., Poisson new call arrivals, exponential channel holding times, exponential cell sojourn times. While the entire multicellular system can be characterised by a Markov process with many dimensions and a complex state space, obtaining performance measures directly from this characterisation is an intractable problem. Our approximate analysis approach is an extension of the iterative technique that has been used in the past for macrocellular networks (see, for example, [8], [16], [9], [5]). Each cell is analysed in isolation, assuming Poisson processes for handoff arrivals into the cell. Blocking probabilities from this analysis yield handoff arrival rates for the next iteration. These iterations are continued until an appropriate convergence criterion is met. The main effort in adapting this standard approach to our problem is the isolated macrocell analysis, especially when overflows, repacking, and mobility changes are introduced. We develop approximations for these analyses, and show that the numerical results obtained compare favourably with those obtained from a detailed simulation. Whereas the analysis is based on iterative calculations on an isolated cell, the simulation is of a multimacrocell system, and actually simulates call handovers between cells, slow call overflow and repacking between macrocells and microcells, and mobility changes.

Related work on this problem has appeared in [21], [10], [9], [22], [4], [13], [5]. In [21] a cellular system model with call overflow and repacking between two layers of overlapping cells is considered. There are no call mobility considerations in this paper. The technique is based on the observation that, with repacking, the underlying Markov chain is equivalent to that of a certain circuit switched network. The Erlang fixed point approach is used to approximately analyse this network. The approach, however, leads to a number of "link" constraints that is exponential in the number of cells. The accuracy of the results is found to vary from 15% to 40% depending on the number of channels. In [9] a hierarchical model with three layers is considered; there are two call classes, and calls can overflow to higher layers. Overflow processes are modeled as interrupted Poisson processes (IPP) and are not repacked. Mobility changes are not considered, and no simulation results are provided. In [4], three types of calls are considered in a single cell with a two-tier architecture. The types of calls are classified on the basis of their access to the different tiers. The model does not include handovers, or repacking. In [5] and [10] a nonhomogenous system (cell sizes are different, arrival rates vary from cell to cell, arbitrary routing between cells, and a general overlap structure between layers) is analysed by iterating all the cells together. In [5] the overflow processes between layers are modelled by using two moments, whereas in [10] the composite overflow processes are approximated as Poisson. In [10], calls are identified as being fast or slow depending on their sojourn time in a cell; a call identified as fast is handed over to a higher layer macrocell. These papers do not consider repacking, and only analytical approximations are presented without validating simulation results.

In [22], a procedure for identifying the mobility class of a call (i.e., fast call or slow call) is proposed. A mobile determines its mobility based on its microcell sojourn time. This information is used to determine the base station (at the macrocell or at the microcell) which will handle the call during origination or handoff of the call. A similar approach for identifying fast calls is proposed in [13], and in addition analysis of grade-of-service is done for a two layer system. The latter paper, however, does not consider slow call repacking and mobility changes; also only analytical results are presented.

The remainder of the paper is organised as follows. In Section II we describe the model, list the notation used, and define the performance measures. An approximate analysis for this model is developed in Section III. In Section IV we provide numerical results that show how accurate the analysis is in comparison with simulations of the model. The conclusions and an outline of further work are presented in Section V.

II. The Model, Notation and Terminology

A. Handovers, Repacking and Signalling

We define a handoff (or handover) as any event that causes the system to seek a new channel for an existing call in the system. Handoffs occur due to cell boundary crossings (i.e., a "*radio–reason*" handoff), mobility changes, or repacking. A radio–reason handoff occurs whenever a slow call crosses a microcell boundary, or a fast call crosses a macrocell boundary.

When a fast call changes mobility to become a slow call, an attempt is made to assign it to a channel in the microcell in which it is located. If this attempt fails, then the call is retained in the macrolayer. When a slow call in the microlayer changes mobility, an attempt is made to assign it to a channel in the macrolayer. If this attempt fails, the call is not retained in the microlayer but is dropped. If this call is retained in the microlayer, it will encounter a large number of cell boundary crossings. This is not desirable since, after adding substantially to the signalling traffic, it is very likely to get dropped anyway. No harm is done by dropping the call provided the overall call dropping probability is better than the operator's promised grade-of-service (say, e.g., 0.1%). Channel reservation for fast calls in the macrolayer can be used to control this dropping probability. We have not considered channel reservation in this paper; but see [19].

If a slow call in the macrolayer moves across a microcell boundary, then an attempt is always made to hand the call over to a microcell channel. If there is no such channel, then the slow call is retained in the macrolayer.

Handovers are also caused by the repacking of slow calls occupying macrolayer channels; i.e., slow calls that are assigned channels in the macrolayer are moved back to the microlayer on availability of channels in their respective microcells. Channels in the macrolayer are thus freed up.

Note that the repacking of a slow call in this way is triggered by a slow call departure from a microcell; a slow call in the macrolayer does not need to constantly monitor the occupancy of its microcell. Thus slow calls are handled in a macrocell only when their corresponding microcell is fully occupied. This increases the capacity of the system, but additional signalling will be incurred due to the channel reassignments.

Channel reassignments and handoffs cause signalling traffic, and hence load the call processing systems. The set of events that contribute to the signalling traffic are new call arrivals, cell boundary crossings, mobility changes and repacking.

B. Model Parameters and Notation

New call arrival processes for the various macrocells are independent Poisson processes. Each arrival into a macrocell is fast or slow with a certain probability. The probability that an arriving call is fast or slow may be different in different macrocells. A call arriving to a macrocell is assumed to be located in a particular microcell within the macrocell with a certain probability. The conversation time for a call, and a mobile's sojourn time in a cell are assumed to be exponentially distributed. Further, the intervals at which a mobile changes its mobility are also assumed to be exponentially distributed. In practice, these intervals will include the time to reliably detect the mobility change.

Macrocells are numbered and are indexed by integers $\{1, 2, ...\}$. There are m_i microcells in the i^{th} macrocell. The microcells in the i^{th} macrocell are numbered using double indices $(i, j), 1 \le j \le m_i$. Define

N_i = the number of channels assigned to macrocell i in the macrolayer

$n_{i,j}$ = the number of channels assigned to microcell j in macrocell i

Λ_i = the total arrival rate of new calls (fast and slow) in macrocell i

ϕ_i = the probability that a new call in macrocell i is a fast call

$\omega_{i,j}$ = the probability that a call originating in macrocell i is physically located in microcell j

μ^{-1} = the mean conversation time of a call in the system; *taken to be 1 always; thus all times are normalised to the mean call duration.*

$\sigma_{i,j}^{-1}$ = the mean sojourn time of a slow call in the microcell (i,j)

Σ_i^{-1} = the mean sojourn time of a fast call in the macrocell i

Γ = the rate of change of mobility of fast calls

γ = the rate of change of mobility of slow calls

The mobility change model is to be understood as follows: a call that is now a slow call, will become a fast call after a random time that is exponentially distributed with mean $1/\gamma$, provided, of course, that the conversation lasts that long. We further define:

$R_{i,k}$ = the probability that a call leaving macrocell i enters macrocell k

$r_{(i,j),(k,l)}$ = the probability that a call leaving microcell (i, j) enters microcell (k, l)

Performance Measures: In this paper we analyse the models to obtain the *new call blocking probability* for each call class (i.e., slow or fast); i.e., the probability that a new call of that class is blocked on arrival to the system. We denote the blocking probabilities by B_{fast} and B_{slow}. Other performance measures of interest would be: handoff blocking probabilities, call dropping probabilities, and the system signalling rate for setting up new calls, and for handling handoffs.

III. Analysis of the Model

A. The Approximate Analysis Approach

There are M cells, indexed by $i \in \{1, 2, \ldots, M\}$, and cell i has m_i microcells, indexed by $j \in \{1, 2, \ldots, m_i\}$. We define the following stochastic processes for $t \ge 0$.

For $1 \le i \le M$, define

$X^{(i)}(t)$ = number of fast calls in the macrolayer of cell i,

$Y^{(i)}(t)$ = number of slow calls in the macrolayer of cell i,

and, for $1 \le j \le m_i$,

$Y_j^{(i)}(t)$ = number of slow calls in the *macrolayer* of cell i that are located (at time t) in *microcell* (i, j); (of course, $Y^{(i)}(t) = \sum_{j=1}^{j=m_i} Y_j^{(i)}(t)$),

$Z_j^{(i)}(t)$ = number of slow calls in the *microlayer* that are located (at time t) in *microcell* (i, j),

and denote by

$\xi^{(i)}(t) = (X^{(i)}(t), ((Y_j^{(i)}(t), Z_j^{(i)}(t)), 1 \le j \le m_i))$.

With our stochastic assumptions (Poisson new call arrivals, exponentially distributed channel holding times, exponentially distributed cell sojourn times, and Markovian call routing between cells), the stochastic process $\{(\xi^{(i)}(t), 1 \le i \le M), t \ge 0\}$ is a Markov process. The number of calls in each layer is restricted by the total number of available channels in that layer. Hence we have a finite state space for this process. For finite and positive values of all the rate parameters, this Markov process is irreducible and hence positive recurrent; thus a stationary distribution exists. In principle, the stationary blocking and dropping probabilities can be obtained from this stationary distribution. Owing to the several special features of this model (handoffs, overflows, repacking, and mobility change), the stationary distribution does not have a "product form". Further, owing to the large size of the state space, direct numerical computation is intractable. Consequently, we resort to an approximate analysis technique similar to the one adopted by several previous researchers in this area (for example, [8], [16]).

The process in the cell i, i.e., $\{\xi^{(i)}(t)\}$ is analysed in isolation, assuming that the arrival process of handoffs from the neighbouring cells is Poisson. This is done for every cell, and, using the intercell routing probabilities, handoff rates between the various cells are obtained. The isolated cell analyses are repeated with these new handoff rates. This iterative process is begun with some initial value of handoff rates entering each cell (e.g., zero rates). If this iterative calculation converges (as it does in all the cases that we have studied) then the limiting probability distribution provided by the iteration at the ith cell is taken to be the stationary distribution of the ith marginal of the process $\{\xi^{(i)}(t), 1 \leq i \leq M\}$. Since new call arrivals are Poisson, this yields an approximation for the new call blocking probability.

In this paper we (i) develop the isolated cell analysis with Poisson arrivals, with macrocells, microcells, repacking and mobility changes, and (ii) examine the accuracy of this approximate analysis procedure for a homogeneous cellular network (i.e., all cells are identical, having the same number of microcells, arrival rate, mean call holding time and sojourn time, and also the same number of channels in the macrolayer and microlayer). Such a homogenous model can be used to model the central cells in a large array of cells in which the nonhomogeneity is only in the boundary cells. Note that the models analysed in [9] and [13] are also homogenous.

B. Additional Notation for the Analysis of an Isolated Cell in the Homogeneous Model

For the homogeneous model, in the stationary regime, we drop the superscript (i) from the various notations. We denote the stationary marginal random variable for $\{X^{(i)}(t)\}$ by X, for $\{Y^{(i)}(t)\}$ by Y, for $\{Y_j^{(i)}(t)\}$ by Y_j, and that for $\{Z_j^{(i)}(t)\}$ by Z_j. Also, for the homogeneous case, the notation in Section II-B yields $N_i = N, n_{i,j} = n, m_i = m, \Lambda_i = \Lambda, \phi_i = \phi, \omega_{i,j} = \frac{1}{m}, \sigma_{i,j} = \sigma, \Sigma_i = \Sigma$.
Define λ_o = arrival rate of new fast calls in a macrocell; these are serviced in the macrolayer (thus, $\lambda_o = \Lambda\phi$); λ_h = arrival rate of handed-off fast calls in the macrolayer; λ_m = arrival rate of fast calls in the macrolayer due to mobility change of slow calls in the microlayer; λ_f = total arrival rate of fast calls in the macrolayer. Hence

$$\lambda_f = \lambda_o + \lambda_h + \lambda_m \tag{1}$$

We also define the following arrival rates of slow calls. ψ_o = arrival rate of new slow calls in a microcell (hence $\psi_o = \frac{\Lambda(1-\phi)}{m}$); ψ_h = arrival rate of slow handoff calls in a microcell; ψ_m = arrival rate of slow calls in a microcell due to change of mobility of fast calls in the macrolayer; ψ = the total arrival rate of slow calls in a microcell. Hence

$$\psi = \psi_o + \psi_h + \psi_m \tag{2}$$

Further, we denote by λ_s the rate of arrival of overflow slow calls to a macrocell. The rates $\lambda_h, \lambda_m, \lambda_s, \psi_h$ and ψ_m are a priori unknown and are calculated iteratively after assuming an initial value for them. The dependence of these rates on the various random variables defined in Section III-A is shown in Section III-B.1.

B.1 Calculation of Various Stationary Rates

The rate at which fast calls handoff from a macrocell is Σ. A handed off call can enter any one of its l neighbours with equal probability. All the cells are assumed to be identical and hence $E(X)$ (see the stationary marginal random variables defined above) is taken as the expected number of fast calls in any cell in the macrolayer. It is clear that in the stationary regime, the arrival rate due to handoffs from a single neighbour cell is $E(X)\Sigma/l$. These arrivals occur from all the l neighbours of a cell. Hence

$$\lambda_h = E(X)\Sigma \tag{3}$$

$E(Y)$ is the expected number of slow calls in the macrolayer, and $E(Z)$ is the expected number of slow calls in a microcell. Assuming homogeneity among the microcells within a cell also, we have $E(Y_j) = \frac{E(Y)}{m}$, $1 \leq j \leq m$. Since slow calls occupying macrolayer channels are always attempted to be handed off to the microlayer when they cross a microcell boundary, we have

$$\psi_h = (E(Z) + \frac{E(Y)}{m})\sigma \tag{4}$$

Also slow calls from any of the m microcells of a macrocell may become fast calls at rate γ. Therefore

$$\lambda_m = mE(Z)\gamma \tag{5}$$

Since all the microcells in a cell are considered to be identical, a fast call in the macrocell is located in any one of the microcells with probability $1/m$. $E(X)\Gamma$ is the rate at which fast calls in the macrolayer generate slow calls due to mobility change. Hence

$$\psi_m = \frac{E(X)}{m}\Gamma \tag{6}$$

$E(X), E(Y)$ and $E(Z)$ are again functions of the net arrival and net service rates of fast and slow calls in a cell. Hence these can be computed iteratively, and then used to compute the blocking probabilities.

C. Analysis of the Isolated Cell Model Without Repacking

In this model, a slow call that arrives in a cell and is served in the macrolayer, owing to the nonavailability of a channel in the microlayer, is retained in the macrolayer until it requires radio-reason handoff, or crosses a microcell boundary, or until it completes the conversation. If the slow call crosses a microcell boundary (even if it is in the same macrocell) then a channel is first sought for it in the microcell that it enters.

The isolated cell model comprises m groups of n servers each, corresponding to the microcells, and one group of N servers corresponding to the macrolayer channels. Slow calls arrive to the microcell $j, 1 \leq j \leq m$, in a Poisson process at the rate ψ; fast calls arrive to the macrolayer channels in a Poisson process at the rate λ_f. A slow call finding its microcell full overflows to the macrocell channels. A fast call holds a macrocell channel for an exponentially distributed duration with rate $\mu + \Sigma$, but changes class to slow at the rate Γ. Similarly, a slow call in the macrolayer holds a channel for an exponentially disributed time with rate $\mu + \sigma$, but changes mobility at the rate γ. Observe that, without mobility changes, this model is just the classical overflow model that arises in telephone trunk engineering. Owing to the large number of microcells, we assume that the overflow process is Poisson. We will show how this approximation works in comparison with simulations. In contrast in [9] the overflow process is modelled by an IPP; for our situation, where we are modelling several new features, considering the additional state of the IPP would further complicate the analysis. Simulations show that our approximations are adequate.

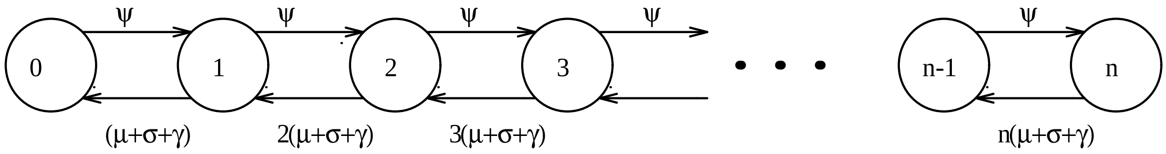

Fig. 1. Transition rate diagram for the microcell process $Z(t)$, with no repacking

C.1 Stationary Analysis of the Microlayer

For a microcell in isolation, assuming Poisson arrival processes, $\{Z(t)\}$ is a Markov chain on $\{0, 1, ..., n\}$, with the transition rate diagram shown in Figure 1; here ψ is as defined in Section III-B.
Hence, the stationary probability, $P(Z = n)$ is calculated from the Erlang formula, $Erlang_B(a, n)$, which is defined as

$$Erlang_B(a, n) = \frac{a^n/n!}{\sum_i a^i/i!}, \quad (7)$$

where a is the offered load in Erlangs, and is given for our model as $a = \frac{\psi}{\mu+\sigma+\gamma}$.
Finally, $E(Z)$ (as needed in Equations 4 and 5) is computed, using Little's theorem, as

$$E(Z) = \frac{\psi}{(\mu + \sigma + \gamma)}(1 - P(Z = n)) \quad (8)$$

C.2 Stationary Analysis of the Macrolayer

Slow calls blocked from microcells, or those changing mobility, arrive into the macrolayer. Hence in the isolated cell model, the process $\{(X(t), Y(t))\}$ depends on the process $\{Z(t)\}$. If the number of microcells in a macrocell is large, then we can expect that the dependence of the macrolayer process on any particular microcell will be small, and also the microcells will be weakly dependent among themselves. With this in mind, we approximate this dependence by using the stationary probabilities obtained for $\{Z(t)\}$, and hence model $\{(X(t), Y(t))\}$ as a Markov chain with state space $S = \{(n_f, n_s); n_f + n_s \leq N\}$.

The macrolayer has new fast call arrivals in a Poisson stream. A fast call can leave the macrolayer for one of three reasons: on call completion, or on cell boundary crossing, or on a mobility change with the probability that the microcell in which it is located has a free channel. To account for this last possibility, we need the conditional probability distribution of $\{Z(t)\}$, conditioned on the states of the process $\{X(t), Y(t)\}$. However, as stated earlier, as an approximation, we use the stationary probabilities of the process $\{Z(t)\}$. Hence the rate at which a fast call leaves the macrolayer due to mobility change is calculated as $\Gamma(1 - P(Z = n))$. A slow call leaves the macrolayer either on call completion or on cell boundary crossing; from the point of view of a single isolated cell model, a slow call, in the macrolayer, that crosses its microcell boundary is seen as leaving the macrolayer (since an attempt is made to serve it in the microlayer of the neighbouring cell; see Section II); actually if the neighbouring microcell is full then the call may be retained in the macrolayer, but this will be viewed as a new overflow arrival from the microlayer in our analysis. Let μ_f denote the total rate at which a fast call leaves a macrocell in the macrolayer, and μ_s denote the total rate at which a slow call leaves the macrolayer. From the arguments above, we have the relations:

$$\mu_f = \mu + \Sigma + \Gamma(1 - P(Z = n)) \quad (9)$$

$$\mu_s = \mu + \sigma \quad (10)$$

Slow calls arrive into the macrolayer when the microcell in which they are located has no free channels. New and handed off slow calls arrive to each microcell at the rate $\psi_o + \psi_h$. Hence the rate of arrival of overflow slow calls to the macrolayer is

$$\lambda_s = m(\psi_o + \psi_h)P(Z = n) \quad (11)$$

The arrival rate of fast calls to the macrolayer, λ_f, is given by Equation 1 and the expressions in Section III-B.1.

A fast call becomes a slow call and is retained in the macrolayer if all the channels in its corresponding microcell in the microlayer are occupied. As above, we assume that a fast call that becomes slow finds its corresponding microcell full with probability $P(Z = n)$. Further, a slow call in the macrolayer retains its channel if it becomes fast. With these observations we define the rates

$$\gamma_{fs} = \Gamma P(Z = n) \quad (12)$$

$$\gamma_{sf} = \gamma \quad (13)$$

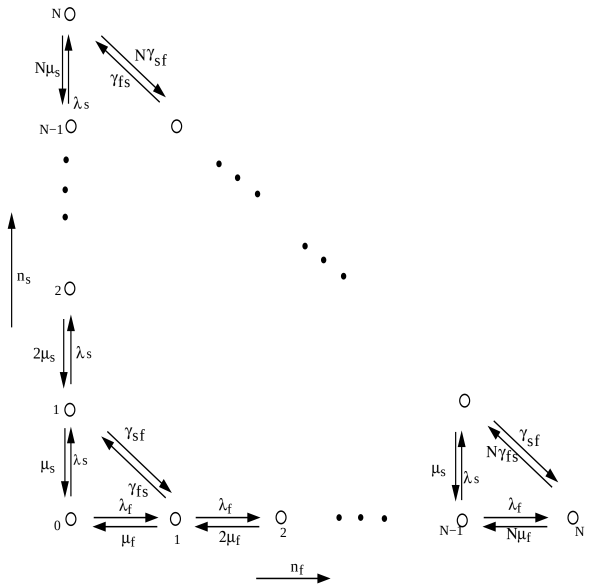

Fig. 2. Transition rate diagram for macrocell process $\{(X(t), Y(t))\}$, with no repacking

It is now clear that, with the assumptions made and the notation defined, $\{(X(t), Y(t))\}$ has the transition diagram shown in Figure 2.

It is easily seen that the transition diagram in Figure 2 is the same as that of the closed Markovian queueing network shown in Figure 3. There are two nodes, 1 and 2; node 1

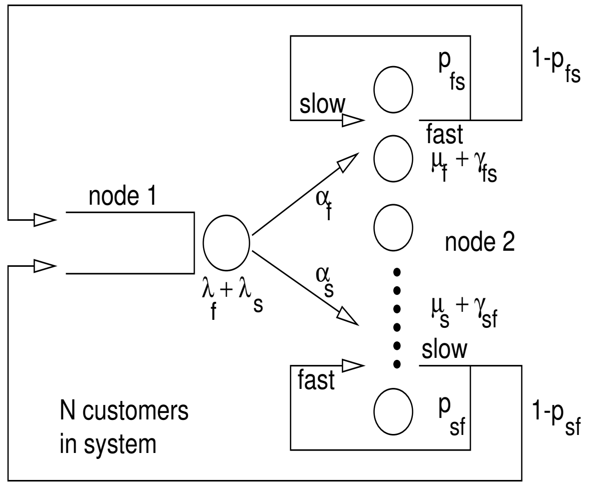

Fig. 3. Equivalent product form queueing network model, for the Markov chain in Figure 2

represents the arrival process and node 2 the service process. There are three types of calls: the incoming calls that are only at node 1, and fast and slow calls that are at node 2. The service rate at node 1 is $\lambda_s + \lambda_f$; customers at node 1 depart as fast or slow calls according to the probabilities α_f and α_s where $\alpha_f = \frac{\lambda_f}{\lambda_f+\lambda_s}$, $\alpha_s = 1 - \alpha_f$. The service rates at node 2 are $\mu_f + \gamma_{fs}$, and $\mu_s + \gamma_{sf}$, for fast and slow calls, respectively. The mobility changes are taken care of by class changes. The value p_{fs} indicates the probability that a fast call leaving node 2 returns to node 2 as a slow call; here, $p_{fs} = \frac{\gamma_{fs}}{\mu_f+\gamma_{fs}}$. Similarly we obtain p_{sf}. We can use the BCMP theorem [6] to show that there is a product form solution for the stationary distribution of the process $\{(X(t), Y(t))\}$. The product form stationary distribution of the random vector (X, Y) is

$$\pi(n_f, n_s) = \frac{(a_1)^{n_f}/n_f! \times (a_2)^{n_s}/n_s!}{\sum_{n_f+n_s \le N} (a_1)^{n_f}/n_f! \times (a_2)^{n_s}/n_s!} \qquad (14)$$

where,

$$a_1 = (\lambda_f + \lambda_s p_{sf})/(\mu_f + (1-p_{sf})\gamma_{fs}) \qquad (15)$$

$$a_2 = (\lambda_s + \lambda_f p_{fs})/(\mu_s + (1-p_{fs})\gamma_{sf}) \qquad (16)$$

Given $\pi(\cdot,\cdot)$, the new call blocking at the macrolayer is given by

$$P(X+Y=N) = \sum_{n_f+n_s=N} \pi(n_f, n_s) \qquad (17)$$

C.3 Calculations from $\pi(n_f, n_s)$

The problem of finding $P(X+Y=N)$ is the same as that of finding the blocking probability in an Erlang-B model in which two classes of customers arrive in Poisson processes; one class brings a load of a_1 Erlangs and the other a load of a_2 Erlangs. We can merge the two Poisson streams into one with a holding time distribution that is the probabilistic mixture of the two, and which brings a load of a_1+a_2. Since the Erlang blocking formula is insensitive to the holding time distributions and depends only on the load, we have (exactly)

$$P(X+Y=N) = \frac{(a_1+a_2)^N/N!}{\sum_{i \le N}(a_1+a_2)^i/i!} \qquad (18)$$

Also from Little's theorem we have:

$$E(X) = a_1(1 - P(X+Y=N)) \qquad (19)$$

$$E(Y) = a_2(1 - P(X+Y=N)) \qquad (20)$$

The expected values $E(Z)$ (given by Equation 8), $E(X)$ and $E(Y)$ are used to calculate the arrival rates due to handoffs and mobility changes in the macrolayer and the microlayer as described in Section III-B.1. With these new rates, the next iteration is performed. Starting with $\lambda_h = 0, \lambda_m = 0, \psi_h = 0, \psi_m = 0$ the iterations are continued until the change in the rates is less than a small $\epsilon > 0$.

When the iterations terminate, the final values for the single isolated cell analysis yield the desired blocking probabilities. The fast call blocking is equal to the probability that the macrocell is full and is given by the Equation 18. The slow call blocking is approximated by the product of macrocell blocking and microcell blocking. Hence

$$B_{fast} = P(X+Y=N) \qquad (21)$$

$$B_{slow} = P(Z=n)P(X+Y=N) \qquad (22)$$

D. Analysis of the Isolated Cell Model with Repacking

Repacking refers to the policy that a slow call using a macrocell channel is shifted to a channel in the microcell in which it is located as soon as one frees up. Thus if a slow call is occupying a macrolayer channel, it *implies* that its microcell is full. Repacking is similar to a handoff from the point of view of the signalling required to achieve it. Hence there is the question of the improvement in the blocking performance due to repacking versus the increase in the signalling load. The "aggressive" repacking strategy that has been described here may not be the best to adopt, as it may cause excessive signalling load without much gain in blocking performance. For a performance study of various "lazy" repacking strategies see [20]. We proceed in this paper with the assumption of aggressive repacking.

We first obtain an approximation to the blocking probabilities when there are no mobility changes, and later include mobility changes. Without mobility change, the isolated cell model comprises m banks of n servers each, corresponding to the microcells, and one bank of N servers corresponding to the macrolayer channels. Slow calls arrive to the microcell $j, 1 \le j \le m$, in a Poisson process at the rate ψ (see Section III-B); fast calls arrive to the macrolayer channels in a Possion process at the rate λ_f. A slow call finding its microcell full overflows to the macrocell channels. When a slow call departs from a microcell, a slow call located in that microcell that is holding a macrocell channel is moved to the vacated microcell channel. A slow call holds a channel (microcell or macrocell) for an exponentially distributed time with rate $\mu + \sigma$. A fast call holds a macrocell channel for an exponentially distributed duration with rate $\mu + \Sigma$. Define $X_j(t) = Y_j(t) + Z_j(t)$ for $1 \le j \le m$; i.e., $X_j(t)$ is the total number of slow calls in microcell j at time t. Note that, owing to repacking, $Z_j(t) = \min\{X_j(t), n_j\}$ and $Y_j(t) = X_j(t) - Z_j(t)$. It is clear that the process $\{(X(t), X_j(t), 1 \le j \le m)\}$ has a product form stationary distribution since we have a multiclass resource sharing model with a coordinate convex *partial sharing* policy (see [11]). In principle, the blocking probabilities can be computed from this product form distribution. Since this is a partial sharing policy, Kaufman's recursion does not apply. For the large numbers of channels (order of 100), and the large numbers of microcells that we will consider, direct computation is not tractable. We will use this product form distribution, however, to make certain exact arguments in the approximate analysis that we now develop. In [21], the product form is observed to be equivalent to that of a circuit switched network, and an Erlang fixed point iteration is used; an accuracy of 15% to 40% is reported.

D.1 Stationary Analysis of the Microlayer

As before, we first consider a microcell process $\{Z_j(t)\}$, for some $j, 1 \leq j \leq m$. We model $\{Z_j(t)\}$ approximately as a Markov chain on the state space $\{0, 1, \ldots, n\}$. When $Z_j = n$, the macrolayer holds at least one slow call that belongs to this microcell with probability $P(Y_j > 0/Z_j = n)$. Hence, owing to repacking, the transition rate from the state $Z_j = n$ to $Z_j = n-1$ is $P(Y_j = 0/Z_j = n)n(\mu + \sigma)$. The remaining transition rates are unaffected by repacking and are as in Figure 4. For the purpose of blocking probability calculations in later sections, we need to obtain the conditional probability $P(Z_j = n \mid Y_j = 0)$. Observe that, owing to the fact that slow calls are always offered to the microcell first, and owing to repacking, when the set of states with $Y_j > 0$ is exited then $Z_j = n$; further, the set of states with $Y_j > 0$ is entered only from the set of states with $Z_j = n$. It follows that the process $\{Z_j(t)\}$ conditioned on $Y_j(t) = 0$ is just the Erlang-B process with offered load $\frac{\psi}{\mu+\sigma}$ and number of servers n. Hence

$$P(Z_j = n \mid Y_j = 0) = Erlang_B(\psi/(\mu + \sigma), n) \qquad (23)$$

Observe that we do not have $P(Y_j = 0 \mid Z_j = n)$; hence the analysis of the Markov chain for $Z(t)$ is not possible. We will see, however, that this analysis is not necessary for the calculation of blocking probabilities.

D.2 Stationary Analysis of the Macrolayer

As in Section III-C.2, we analyse the process $\{(X(t), Y_j(t), 1 \leq j \leq m\}$ by approximating its interactions with $\{(Z_j(t), 1 \leq j \leq m)\}$ using stationary probability distributions.

When $Y_j = 0$, a slow call from microcell j is offered to the macrolayer only when it is blocked in the microcell j into which it arrives. This happens with probability $P(Z_j = n \mid Y_j = 0)$, which has been obtained above. When $Y_j > 0$, since repacking is done, the microcell j must be full, and every slow call arrival to this microcell will overflow into the macrolayer. Thus, defining $B = P(Z_j = n \mid Y_j = 0)$, the transition rate from state $(x, y_1, y_2, \ldots, y_j = 0, \ldots, y_m)$ to $(x, y_1, y_2, \ldots, y_j = 1, \ldots, y_m)$ is ψB while the transition rate from states $Y_j = k$ to $Y_j = k+1$ is ψ for $1 \leq k \leq N$.

We now develop an approximate analysis for the process $\{(X(t), Y(t) = \sum_{i=1}^{m} Y_j(t))\}$, which has the state space, $S = \{(n_f, n_s); n_f + n_s \leq N\}$.

Define the random process $K(t) = \sum_{j=1}^{j=m} I_{\{Y_j>0\}}(t)$; where $I_{\{\cdot\}}(t)$ is the indicator process of the set $\{\cdot\}$. Thus, $K(t)$ is the number of microcells that have slow calls in the macrolayer. Observe that, given the process $K(t)$, we can obtain the transition rates for the coordinate $Y(t)$ of the process $\{X(t), Y(t)\}$. When $K(t) = k$ and $Y(t) = n_s$, we see that overflowed slow calls arrive into the macrolayer from k microcells with a total arrival rate $k\psi$, while calls arrive from the remaining microcells with arrival rate $(m-k)B\psi$. Therefore the net arrival rate of slow calls to the macrolayer when $Y(t) = n_s$ and $K(t) = k$ is given by,

$$\lambda_{n_s,k} = k\psi + B(m-k)\psi \qquad (24)$$

Let $\mu_s = (\mu + \sigma)$. When a slow call departs from one of these k microcells, a slow call that belongs to that microcell is moved from the macrolayer to the microlayer. Due to this repacking, the rate of departure of a slow call from the macrolayer when $(Y(t) = n_s, K(t) = k)$ is given by

$$\mu_{n_s,k} = n_s\mu_s + kn\mu_s \qquad (25)$$

This is because a slow call departs from the macrolayer even if one of the n calls in any of the k full microcells departs.

Unless we keep track of $Y_j(t), 1 \leq j \leq m$, we do not know the value of $K(t)$. To obtain an approximate analysis of $\{(X(t), Y(t))\}$, we estimate a value for $K(t)$, given $Y(t)$, and use this estimate in the transition rate formulas shown above. Thus given the number of slow calls in the macrolayer we want to obtain an estimate of the number of microcells they belong to.

We do this by considering an *urn* model with m urns (corresponding to the m microcells), into which $n_s(= Y(t))$ balls are placed in succession as follows; at the end of placing the n_s balls, the number of nonempty urns corresponds to K. The first ball is thrown into any one of the urns with equal probability (this corresponds to the fact that the first slow call to be handled by the macrolayer comes from any of the microcells with equal probability). Now, given that there are exactly j microcells that have at least one slow call in the macrolayer, the rate of arrival, into the macrolayer, of a slow call from any of these j cells is $j\psi$ while the rate from the other cells is $(m-j)B\psi$. Thus the next slow call arrives from these j cells with probability $j/(j + (m-j)B)$. The probability that the next slow call is from the rest of the cells is $(m-j)B/(j + (m-j)B)$. Hence in the urn analogy, if there are j occupied urns, the next ball is thrown in such a way that the number of occupied urns increases by 1 with probability $(m-j)B/(j + (m-j)B)$. Note that, since B will be small, the next ball is thrown into occupied urns with a much larger probability than the unoccupied urns.

Let $p_j^{(i)}$ be the probability that there are j non-empty urns after i balls are thrown into the urns in the manner described. These can be recursively calculated with the following equations ($\delta_{(\cdot)}$ are indicator functions)

$$\begin{aligned} p_j^{(i)} &= p_{(j-1)}^{(i-1)} \frac{(m-j+1)B}{(j-1+(m-j+1)B)} \delta_{j \leq i} \delta_{j \leq m} + \\ & \quad p_j^{(i-1)} \frac{j}{(j+(m-j)B)} \delta_{j<i} \delta_{j \leq m} \end{aligned} \qquad (26)$$

We know that $p_1^{(1)} = 1$ and $p_0^{(0)} = 1$.

Finally, given $Y(t) = n_s$, we estimate the value of $K(t)$ as the expected number of non-empty urns in the above

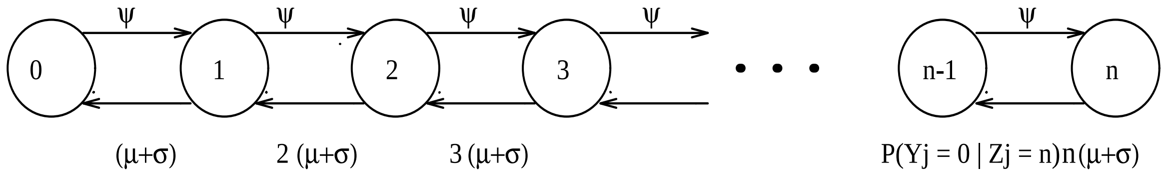

Fig. 4. Transition rate diagram for the microcell process $\{Z(t)\}$, with repacking (mobility change not considered).

urn experiment; i.e., we define

$$k_{n_s} = \sum_{j=1}^{j=m} p_j^{(n_s)} j \tag{27}$$

and use this value for k in Equations 24 and 25, given $Y(t) = n_s$, to obtain the transition rates for the $Y(t)$ coordinate of the process $\{(X(t), Y(t))\}$. Equation 27 requires the computation of $p_j^{(i)}$ for all possible values of i and j. This can be avoided by using a recursion for directly computing k_{n_s} (this is provided in [17]). Thus, we have approximated the process $\{(X(t), Y(t))\}$ by a Markov chain with the transition rates shown in Figure 5; here $\mu_f = \mu + \Sigma$.

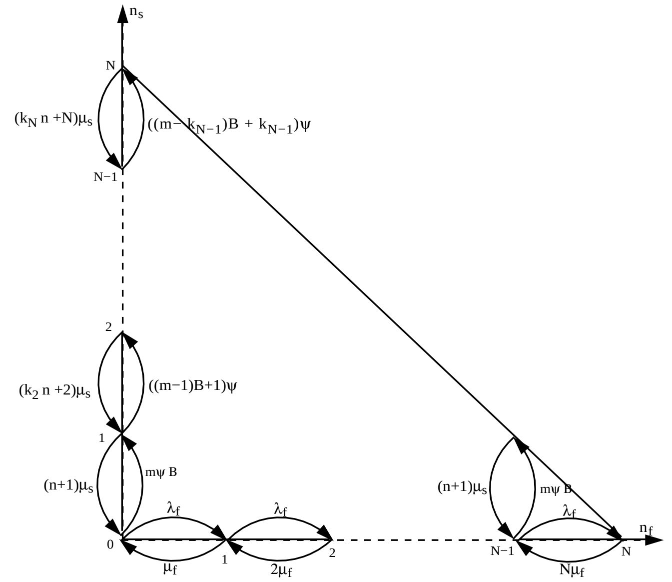

Fig. 5. Transition rate diagram for the macrocell process $\{(X(t), Y(t))\}$, with repacking

Observe that we have a two class blocking model in which the arrival rates and the service rates of each class depend only on the marginal number in that class. Hence, the stationary distribution has the following product form (see [12]). For $(n_f, n_s) \in \{x, y : x \geq 0, y \geq 0, x + y \leq N\}$

$$\pi(n_f, n_s) = G^{-1} g(n_f, n_s) \tag{28}$$

where

$$g(n_f, n_s) = (\lambda_f / \mu_f)^{n_f} / n_f! \times \prod_{i=1}^{i=n_s} \lambda_{i-1} / \mu_i \tag{29}$$

. λ_{n_s} and μ_{n_s} are as given in Equations 24 and 25, with $k = k_{n_s}$. G is the normalisation constant given as

$$G = \sum_{(n_f, n_s)} g(n_f, n_s) \tag{30}$$

The probability that all the channels are occupied in the macrolayer is thus

$$P(X + Y = N) = \sum_{n_f + n_s = N} \pi(n_f, n_s) \tag{31}$$

D.3 Blocking Probabilities for Fast and Slow Calls

Iterative application of the above analysis provides an approximation to the stationary probability distribution for the process $\{(X(t), Y(t))\}$. The blocking probability for fast calls is the probability that all the channels in the macrolayer are occupied. Hence

$$B_{fast} = P(X + Y = N) \tag{32}$$

Slow calls are blocked if the microcell to which they arrive and the macrolayer channels are full. Since all the microcells are considered to be identical, the blocking probability for slow calls arriving into any microcell is

$$B_{slow} = P(Z_j = n, X + Y = N) \tag{33}$$

Writing out the right hand side of Equation 33, we get

$$\begin{aligned} & P(Z_j = n, X + Y = N) = \\ & \quad P(Z_j = n, Y_j = 0, X + Y = N) + \\ & \quad P(Z_j = n, Y_j > 0, X + Y = N) \\ & = \sum_{n_s=0}^{n_s=N} \{ P(Z_j = n, Y_j = 0, Y = n_s, X = N - n_s) + \\ & \quad P(Z_j = n, Y_j > 0, Y = n_s, X = N - n_s) \} \end{aligned} \tag{34}$$

The product form for the stationary distribution of $\{(X(t), Y_j(t) + Z_j(t), 1 \leq j \leq m)\}$ can now be used to establish certain conditional independences (shown in Appendix A, Lemma A.1). These are used to yield the following simplifications.

$$\begin{aligned} & P(Z_j = n, Y_j = 0 \mid Y = n_s, X = N - n_s) = \\ & \quad P(Z_j = n, Y_j = 0 \mid Y = n_s) \\ & = P(Z_j = n \mid Y_j = 0, Y = n_s) P(Y_j = 0 \mid Y = n_s) \\ & = P(Z_j = n \mid Y_j = 0) P(Y_j = 0 \mid Y = n_s) \end{aligned} \tag{35}$$

Similarly

$$\begin{aligned} & P(Z_j = n, Y_j > 0 \mid Y = n_s, X = N - n_s) = \\ & \quad P(Z_j = n \mid Y_j > 0, Y = n_s) P(Y_j > 0 \mid Y = n_s) \end{aligned} \tag{36}$$

Hence, we get

$$\begin{aligned} B_{slow} = & \sum_{n_s=0}^{n_s=N} P(Y = n_s, X = N - n_s) \cdot \\ & \{ P(Z_j = n \mid Y_j = 0) P(Y_j = 0 \mid Y = n_s) + \\ & P(Z_j = n \mid Y_j > 0, Y = n_s) P(Y_j > 0 \mid Y = n_s) \} \end{aligned} \tag{37}$$

$P(Y = n_s, X = N - n_s)$ is obtained from the stationary distribution Equation 28 (see [17]). Also $P(Z_j = n / Y_j = 0) = Erlang_B(\psi/(\mu + \sigma), n)$ as shown in Section III-D.1. Obviously, $P(Z_j = n \mid Y_j > 0, Y = n_s) = 1$, owing to the repacking policy. Finally, recalling the definition of the process $K(t)$, from Section III-D.2, and letting K denote the stationary random variable for $K(t)$, we have

$$P(Y_j > 0 \mid Y = n_s) = \frac{1}{m} E(K \mid Y = n_s) \tag{38}$$

and, of course,

$$P(Y_j = 0 \mid Y = n_s) = 1 - P(Y_j > 0 \mid Y = n_s) \quad (39)$$

We approximate $E(K \mid Y = n_s)$ with k_{n_s} from Equation 27.

Hence we have all the ingredients to compute B_{slow} from Equation 37.

D.4 Including Mobility Changes in the Analysis with Repacking

When there are mobility changes in the system, the model for the macrolayer has a transition structure with "diagonal" transitions due to change of type of calls. Since the transition rates in the $Y(t)$ coordinate are not simply proportional to n_s, the product form distribution now fails to hold. Further, we do not have an analysis of the microcell process $Z(t)$; hence we cannot obtain arrival rates of slow calls due to mobility change as we did in the case with no repacking.

We develop an approximation for the blocking probabilities by viewing a change of mobility as an arrival of a call of the other type. Each slow call arrival is viewed as an arrival of a slow call, and also an arrival of a fast call with the probability that the slow call will change class before it terminates or leaves the microcell in which it arrived. The channel holding times of these two arrivals are adjusted so that the total offered Erlang load due to slow calls remains unchanged. The same is done for fast calls.

Let $\mu_s = \mu + \sigma$ and $\mu_f = \mu + \Sigma$. Define $p_{sf} = \gamma/(\mu_s + \gamma)$; p_{sf} is the probability that a slow call changes mobility before it terminates or hands-off. Thus if λ_f is the total arrival rate of fast calls in a macrocell and ψ that of slow calls in a microcell, then, after including mobility change, the net arrival rate into the fast call stream is taken as λ_1, where

$$\lambda_1 = \lambda_f + m\psi p_{sf} \quad (40)$$

Similarly if λ_2 is the net arrival rate of slow calls into the m microcells in a macrocell, then

$$\lambda_2 = m\psi + \lambda_f p_{fs} \quad (41)$$

where $p_{fs} = \Gamma/(\mu_f + \Gamma)$. We now obtain the modified channel holding rates. Let x_f denote the mean duration of stay of the fast part of a call in the macrocell. It stays for at least a mean duration of $1/(\mu_f + \Gamma)$. Then, with probability p_{fs} it becomes a slow call which again becomes a fast call with probability p_{sf}. Hence with the probability $p_{fs}p_{sf}$, the fast call returns to the system as a fast call and takes an additional duration x_f to leave the system. Hence

$$x_f = 1/(\mu_f + \Gamma) + p_{fs}p_{sf} \times x_f \quad (42)$$

Therefore

$$x_f = \frac{1}{\mu_f + \Gamma(1 - p_{sf})} \quad (43)$$

Also $\mu_1 = 1/x$. From this we obtain

$$\mu_1 = \mu_f + (1 - p_{sf})\Gamma \quad (44)$$

A similar expression can be obtained for the mean duration of stay of a slow call (as a slow call) in a microcell, finally yielding

$$\mu_2 = \mu_s + (1 - p_{fs})\gamma \quad (45)$$

Now λ_1 and λ_2 can be considered to be the net arrival rates of fast and slow calls respectively in a macrocell. Similarly μ_1 and μ_2 can be considered as the net termination rates of the calls in the macrocell. Thus the macrolayer model with mobility changes is analysed by replacing, in the previous analysis, ψ with λ_2/m, λ_f with λ_1, and μ_f, μ_s with μ_1, μ_2.

IV. Comparison of Analysis and Simulation

Recall that our analysis approach involves two levels of approximations. The isolated cell analysis is approximate, even for Poisson arrivals and exponential service times. The multicell analysis is approximate because the handover processes are modelled approximately as Poisson processes, with rates determined from the stationary analysis of the isolated cells. It is important to understand the contribution of the errors in the numerical results from each of these major approximation steps. In Section IV-B we show numerical results obtained from the analysis of a *single* macrocell in isolation. We compare these analytical results with those from a single macrocell simulation with Poisson call arrivals. Overflow and repacking within the macrocell are modelled; the number of microcells and the channel partitioning is varied. These results serve to validate the approximations used in the analysis that we have developed for an isolated cell.

In Section IV-C we show numerical results obtained from the analysis and simulation of a multicell system, for varying arrival rates, and mobility parameters. Whereas the analysis is just iterative calculations on a single macrocell, the multicell simulation models a system of 64 macrocells each with a number of microcells. In the simulation, the assumptions of Poisson new call arrivals, exponential channel holding times, exponential cell sojourn times, and exponential time interval between mobility changes are identical to those in the analysis. However, call mobility, handoffs to neighbouring cells, repacking of slow calls, overflow and mobility changes are all actually simulated. In the simulation, for example, handover calls are routed to neighbouring cells in each layer, and are then handled in the neighbouring cells; when slow calls located in a microcell are in the macrolayer and a slow call departs from that microcell, then a slow call from the macrolayer is repacked; if a slow call in the macrolayer moves to a new microcell with a free channel, the slow call is repacked; etc. Thus the details of the movement and state changes of the calls are simulated exactly as they would be in the full multicell Markov process $\xi^{(i)}(t) = (X^{(i)}(t), ((Y_j^{(i)}(t), Z_j^{(i)}(t)), 1 \le j \le m_i))$ (see Section III).

A. System Parameters for the Numerical Results

The number of channels allocated to each (macro)cell is 80; with a reuse factor of 3 between the macrocells, this would mean that there are 240 channels available in the system. Non-overlapping channel sets are assigned to the macrolayer and the microlayer. A reuse factor of 4 is assumed in the microlayer; hence the set of channels allocated to the microlayer is partitioned into 4 sets. It follows that $N + 4n = 80$. When a macrocell is divided into m microcells, the area of the microcell is $1/m$ times the area of the macrocell. Hence the linear distance that a mobile travels to leave a microcell is $1/\sqrt{(m)}$ times the linear distance the same mobile travels to leave a macrocell. Assuming that fast mobiles are 5 times as fast as the slow mobiles, the sojourn rates of the fast and slow calls (in macrocells and microcells, respectively) are related by $\sigma = \Sigma\sqrt{m}/5$. We also take the mobility change parameters to be related by $\Gamma = 5\gamma$.

Since the value of μ^{-1} (the mean conversation time) is taken as 1, the values of the cell sojourn rates and the rates of change of mobility are normalised to the mean conversation time. Thus, for example, Σ is the average number of macrocells that a fast call crosses during its conversation time.

B. Validation of the Isolated Cell Analysis

TABLE I
SINGLE MACROCELL *without* REPACKING; $\Sigma = 0$; $\Gamma = 0$; $m = 16$; $N = 56$; $n = 6$

Load (erlangs)	Analysis		Simulation	
	B_{fast}	B_{slow}	B_{fast}	B_{slow}
150	0.3347	0.0797	0.3323	0.0849
140	0.2720	0.0572	0.2703	0.0631
130	0.2028	0.0370	0.2006	0.0397
120	0.1307	0.0201	0.1288	0.0224
110	0.0649	0.0082	0.0673	0.0098

TABLE II
SINGLE MACROCELL *with* REPACKING; $\Sigma = 0$; $\Gamma = 0$; $m = 16$; $N = 56$; $n = 6$

Load (erlangs)	Analysis		Simulation	
	B_{fast}	B_{slow}	B_{fast}	B_{slow}
150	0.2268	0.1055	0.2316	0.0984
140	0.1754	0.0757	0.1771	0.0696
130	0.1215	0.0480	0.1246	0.0433
120	0.0710	0.0253	0.0680	0.0228
110	0.0303	0.0093	0.0305	0.0085

TABLE III
SINGLE MACROCELL *without* REPACKING; $\Sigma = 0$; $\Gamma = 0$; $m = 36$; $N = 64$; $n = 4$

Load (erlangs)	Analysis		Simulation	
	B_{fast}	B_{slow}	B_{fast}	B_{slow}
150	0.1790	0.0268	0.1856	0.0273
140	0.1204	0.0158	0.1291	0.0177
130	0.0667	0.0075	0.0688	0.0096
120	0.0268	0.0025	0.0305	0.0028
115	0.0065	0.0005	0.0082	0.0008

Tables I, II, III, and IV show slow call blocking and fast call blocking versus Erlang offered load, in an isolated cell; results are shown from our approximate analysis and from a simulation of the isolated cell model. The specific parameters are given in the figure captions; in each case, the fraction of arrivals that are fast calls is 0.4. Results are shown with and without slow call repacking. Observe that the analysis, in spite of the many approximations made, is quite accurate.

TABLE IV
SINGLE MACROCELL *with* REPACKING; $\Sigma = 0$; $\Gamma = 0$; $m = 36$; $N = 64$; $n = 4$

Load (erlangs)	Analysis		Simulation	
	B_{fast}	B_{slow}	B_{fast}	B_{slow}
150	0.0951	0.0262	0.1041	0.0257
140	0.0564	0.0140	0.0645	0.0156
130	0.0268	0.0058	0.0328	0.0063
120	0.0092	0.0017	0.0104	0.0016
115	0.0045	0.0007	0.0051	0.0007

Owing to the fact that slow calls can use macrocell channels, their blocking probability is much smaller than that of fast calls. This discrepancy, which will result in inefficient system sizing, is corrected by channel reservation (for an analysis with reservation see [19]).

Note that repacking of slow calls will always help to reduce the blocking probability of fast calls, but may increase or decrease the blocking probability of slow calls. The blocking probability can increase since without repacking new slow calls can use free channels in the microlayer, which would have been filled up by repacking. Blocking is a bursty phenomenon, however; when there are free channels in a microcell, and slow calls from that microcell occupy macrolayer channels, then repacking these calls helps slow calls arriving at *other* microcells that may be experiencing a period of blocking. The latter effect is expected to predominate when the number of microcells is large, and the blocking probability of slow calls at their microcells is large; in this scenario there is a large probability that at any time some microcell is overflowing. Observe that for $m = 16, N = 56, n = 6$, introducing repacking increases the blocking of slow calls but substantially reduces that of fast calls. For $m = 36, N = 64, n = 4$, however, there is a slight decrease in slow call blocking when repacking is introduced.

C. Analysis and Simulation Results for the Multicell Model

A multiple macrocell system is analysed using our iterative analysis, and using a multicell simulation; graphs between the Erlang load and the blocking probability are plotted for the parameter values $m = 16, N = 60, n = 5$, and $\psi = 0.4$. The simulation is done for a homogeneous system with 64 macrocells.

Figures 6, 7, 8, and 9, show the results without slow call repacking. Although, done for the multicell case, since $\sigma = \Sigma = 0$, Figure 6 is just another case of the single cell results presented in the previous section; we provide this figure for comparison with the results for the same system parameters with mobility. In Figures 7 and 8, there is

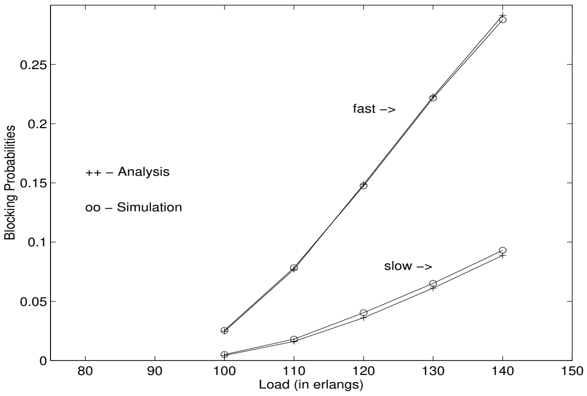

Fig. 6. Multicell system without repacking; $\Sigma = 0; \Gamma = 0$

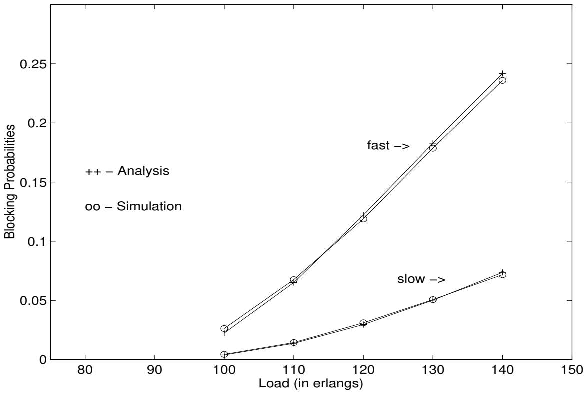

Fig. 7. Multicell system without repacking; $\Sigma = 0.5; \Gamma = 0$

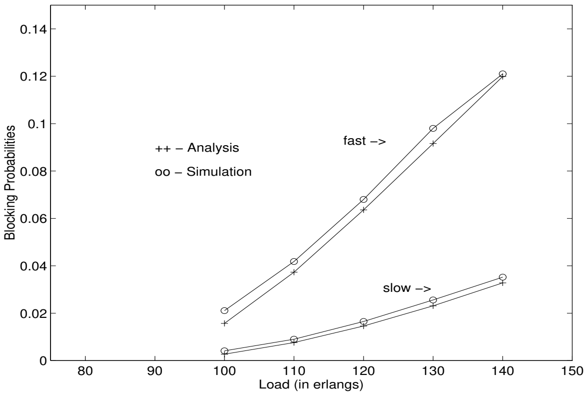

Fig. 8. Multicell system without repacking; $\Sigma = 3.0; \Gamma = 0$

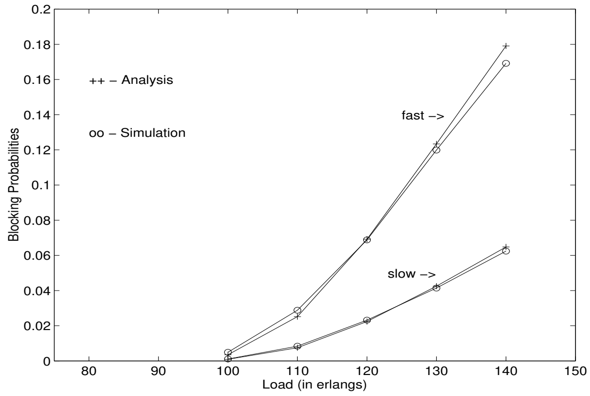

Fig. 9. Multicell system without repacking; $\Sigma = 0.5; \Gamma = 2.0$

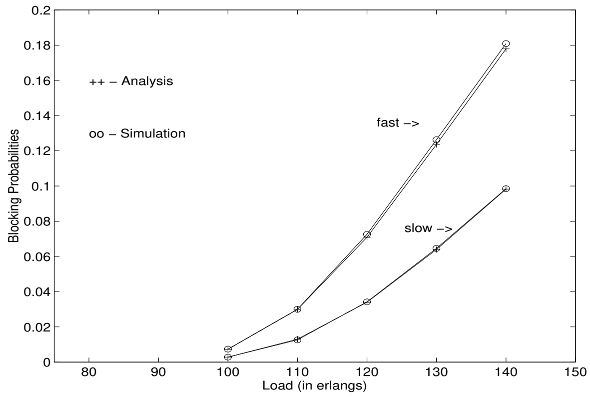

Fig. 10. Multicell system with repacking; $\Sigma = 0; \Gamma = 0$

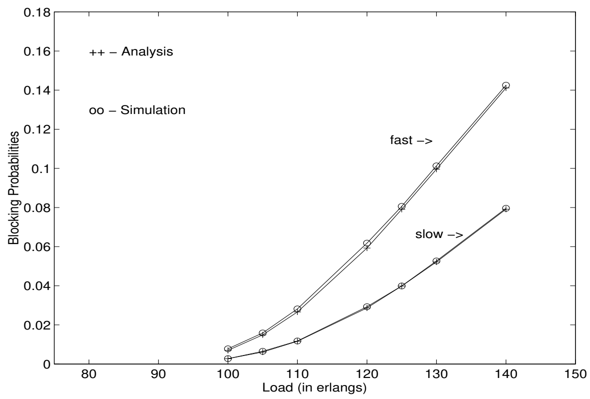

Fig. 11. Multicell system with repacking; $\Sigma = 0.5;\Gamma = 0$

mobility, but no mobility change (in Figure 7, $\Sigma = 0.5$; hence with $m = 16$, the value of $\sigma = 0.4$; in Figure 8, $\Sigma = 3$ and $\sigma = 2.4$). In Figure 9 there is mobility and mobility change; here $\Gamma = 2, \gamma = 0.4$.

In Figures 10, 11, 12, and 13 we provide results with repacking of slow calls. Each of the figures for the cases without repacking has a corresponding figure here, and the results between these should be compared.

Observe, first of all, that the analysis results compare well with those obtained from the simulation. Analysis has the major advantage of requiring just a few minutes of computation time, versus the several hours required for accurate simulation. Thus such an analysis can be very useful in an iterative system sizing process.

In each case, introducing repacking substantially reduces fast call blocking, and increases or slightly reduces the blocking probability of slow calls. Since we do not have channel reservation in these results, slow call blocking is much lower than fast call blocking. Increasing the mobility rate is seen to reduce the blocking probability; this is because with increasing mobility some calls are dropped before they complete conversation, thus reducing the overall occupancy of the channels. We are not controlling dropping probability in these results, as our objective here is only to validate the analysis against simulations.

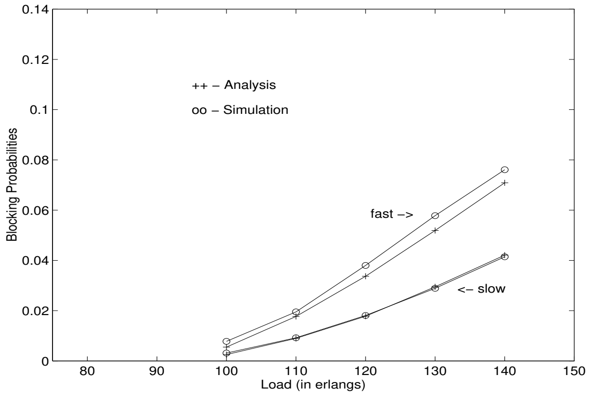

Fig. 12. Multicell system with repacking; $\Sigma = 3.0; \Gamma = 0$

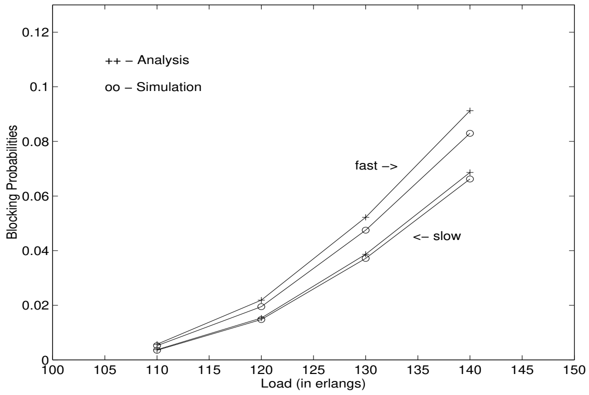

Fig. 13. Multicell system with repacking; $\Sigma = 0.5; \Gamma = 2.0$

V. Conclusion

We have studied the performance of microcellisation in a cellular network, in which the mobiles can be classified as fast or slow. We have developed approximate analyses for calculating the slow and fast call blocking probabilities, and have validated the analyses against a detailed multicell simulation. The approximate analysis is an iterative procedure that utilises an analysis of an isolated cell. We find that in spite of the many approximations made, the analysis results compare well with the simulations. For a large number of microcells, exact analysis of even the isolated macrocell processes is intractable; we are able to obtain approximations that require the analysis of no more than 2-dimensional Markov chains. Such analyses are useful in an iterative procedure for sizing a cellular system to achieve a desired grade of service, since their computation time is much smaller than that for simulations.

It is a relatively straight-forward matter to obtain approximations for dropping probabilities and signalling rates from the analysis [17]. Our analysis procedure in this paper does not permit reservations for fast calls or handovers; the isolated cell analysis needs to be enhanced to accommodate this feature. It is also interesting to explore "lazy" repacking policies. The latter two issues have been addressed in our more recent work reported in [19] and [20]. In these references we have also studied the use of these analytical techniques for system design, i.e., choice of the number of microcells and channel partitioning. More efficient policies for channel allocation to the macrolayer and the microlayer need to be explored.

Appendix

I. Appendix: Proofs of certain conditional independence relations for the analysis with repacking

As observed above in the main text (Section III-D), the process $\{X(t), X_j(t), 1 \leq j \leq m\}$ has a product form stationary distribution. The stationary probability of $\{X(t) = x_0, X_1(t) = x_1, X_2(t) = x_2, \ldots, X_m(t) = x_m\}$ is of the form

$$\Pi(x_0, x_1, x_2, \ldots, x_m) = \frac{1}{G}(\phi_0(x_0)\phi_1(x_1)\phi_2(x_2)\ldots\phi_m(x_m)) \qquad \text{(A.1)}$$

where G is a normalisation constant, and $\underline{x} = (x_0, x_1, x_2, \ldots, x_m)$ is in the state space $\mathcal{S} = \{\underline{x} : 0 \leq x_0 \leq N, 0 \leq x_j \leq n_j + N, 1 \leq j \leq m, x_0 + \sum_{j=1}^{m} x_j \leq N + \sum_{j=1}^{m} n_j\}$. Let $(X, X_1, \ldots, X_m)$ denote the stationary random vector for the process $\{X(t), X_j(t), 1 \leq j \leq m\}$.

Let $S \subset \{0, 1, 2, \ldots, n_j + N\}$ be a set of values that X_j can take.

Lemma A.1:

$$P(X_j \in S \mid \sum_{i=1}^{m} Y_i = y, X = x) = P(X_j \in S \mid \sum_{i=1}^{m} Y_i = y) \qquad \text{(A.2)}$$

Proof: From the product form distribution, we have

$$P(X_j \in S, \sum_{i=1}^{m} Y_i = y, X = x) = \frac{1}{G} \sum_{\{\underline{x}: x_j \in S, \sum_{i=1}^{m} y_i = y, x_0 = x\}} \phi_0(x_0)\phi_1(x_1)\ldots\phi_m(x_m)$$

$$P(\sum_{i=1}^{m} Y_i = y, X = x) = \frac{1}{G} \sum_{\{\underline{x}: \sum_{i=1}^{m} y_i = y, x_0 = x\}} \phi_0(x_0)\phi_1(x_1)\ldots\phi_m(x_m)$$

We denote the set $\{\underline{x}^{(0)} = (x_1, x_2, \ldots, x_m) : (x_0 = x, x_1, x_2, \ldots, x_m) \in \mathcal{S}, x_j \in S, \sum_{i=1}^{m} y_i = y\}$ by $A(x)$ and the set $\{\underline{x}^{(0)} = (x_1, x_2, \ldots, x_m) : (x_0 = x, x_1, x_2, \ldots, x_m) \in \mathcal{S}, \sum_{i=1}^{m} y_i = y\}$ by $B(x)$. Observe that $A(x)$ and $B(x)$ do not depend on x. Denoting these sets by A and B, we have

$$\frac{P(X_j \in S, \sum_{i=1}^{m} Y_i = y, X = x)}{P(\sum_{i=1}^{m} Y_i = y, X = x)} = \frac{\sum_{\{\underline{x}^{(0)} \in A\}} \phi_1(x_1)\ldots\phi_m(x_m)}{\sum_{\{\underline{x}^{(0)} \in B\}} \phi_1(x_1)\ldots\phi_m(x_m)} \qquad \text{(A.3)}$$

We now obtain an expression for $P(X_j \in S \mid \sum_{i=1}^{m} Y_i = y)$. Recall that $\underline{x} = (x_0, x_1, x_2, \ldots, x_m)$. We have

$$P(X_j \in S, \sum_{i=1}^{m} Y_i = y) / P(\sum_{i=1}^{m} Y_i = y) = \frac{\sum_{\{\underline{x}: x_j \in S, \sum_{i=1}^{m} y_i = y\}} \phi_0(x_0)\phi_1(x_1)\ldots\phi_m(x_m)}{\sum_{\{\underline{x}: \sum_{i=1}^{m} y_i = y\}} \phi_0(x_0)\phi_1(x_1)\ldots\phi_m(x_m)}$$

From the definition of the sets $A(x)$ and $B(x)$ above, we have,

$$P(X_j \in S, \sum_{i=1}^{m} Y_i = y)/P(\sum_{i=1}^{m} Y_i = y) = \frac{\sum_{x_0=0}^{N-y} \left(\phi_0(x_0) \sum_{\underline{x}^{(0)} \in A(x_0)} \phi_1(x_1) \dots \phi_m(x_m)\right)}{\sum_{x_0=0}^{N-y} \left(\phi_0(x_0) \sum_{\underline{x}^{(0)} \in B(x_0)} \phi_1(x_1) \dots \phi_m(x_m)\right)} \tag{A.4}$$

Since $A(x)$ and $B(x)$ do not depend on x, Equation A.4 yields

$$P(X_j \in S, \sum_{i=1}^{m} Y_i = y) \mid P(\sum_{i=1}^{m} Y_i = y) = \frac{\sum_{\underline{x}^{(0)} \in A} \phi_1(x_1) \dots \phi_m(x_m)}{\sum_{\underline{x}^{(0)} \in B} \phi_1(x_1) \dots \phi_m(x_m)} \tag{A.5}$$

From Equations A.3 and A.5 we have the result. □

It follows that

$$P(Z_j = n, Y_j = 0 \mid Y = n_s, X = N - n_s) = P(Z_j = n, Y_j = 0 \mid Y = n_s)$$

Similarly, it can also be shown that

$$P(Z_j = n, Y_j > 0 \mid Y = n_s, X = N - n_s) = P(Z_j = n, Y_j > 0 \mid Y = n_s)$$

References

[1] Chih-Lin I., Larry J. Greenstein and Richard D. Gitlin, "A Microcell/Macrocell Cellular Architecture for Low and High Mobility Wireless Users", *GLOBECOM 91* , pp 1006-1011, 1991.

[2] Leonard J. Cimini, Gerard J. Foschini and Chih-Lin I, "Call Blocking and Performance of Distributed Algorithms for Dynamic Channel Allocation in Microcells", *ICC 1992*, pp 1327-1330.

[3] S. A. El-Dolil, W-C. Wong and R. Steele, "Teletraffic performance of Highway Microcells with Overlay Macrocell", *IEEE J. Selected Areas in Communications* , Vol. 7, pp. 71-78, January 1989.

[4] Paul Fitzpatrick, "Performance Analysis of a Layered Network Serving Different User Classes", *IEEE Vehicular Technology Conference*, pp. 431-435, 1996.

[5] Paul Fitzpatrick, Cheng Siong Lee, and Bob Warfield, "Teletraffic Performance of Mobile Radio Networks with Hierarchical Cells and Overflow", *IEEE Journal on Selected Areas in Communications*, Vol. 15, pp. 1549-1557, October 1997.

[6] B. Gelenbe and I. Mitrani, *Analysis and Synthesis of Computer Systems*, John Wiley, 19–.

[7] David J. Goodman, Gregory P. Pollini and Kathleen S. Meier-Hellstern, "Network Control for Wireless Communications", *IEEE Communications Magazine*, December 1992.

[8] Daehyoung Hong and Stephen S. Rappaport, "Traffic Model and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Nonprioritized Handoff Procedures", *IEEE Transactions on Vehicular Technology*, Vol VT-35, August 1986.

[9] Lon-Rong Hu and Stephen S. Rappaport, "Personal Communications Systems Using Multiple Hierarchical Cellular Overlays", *IEEE International Conference on Universal Personal Communications*, pp. 397-401, 1994.

[10] William M. Jolley and Robert E. Warfield, "Modelling and Analysis of Layered Cellular Mobile Networks," *Proc. 13th International Teletraffic Congress, ITC-13*, 1991.

[11] Joseph S. Kaufman, "Blocking in a Shared Resource Environment",*IEEE Transactions on Communication*, Vol. 29, pp 1474-1481, October 1981.

[12] F.P. Kelly, *Reversibility and Stochastic Networks*, 1979.

[13] X.Lagrange and P.Godlewski, "Performance of a Hierarchical Cellular Network with Mobility-Dependent Hand-over strategies", *IEEE Vehicular Technology Conference*, pp. 1868-1872, 1996.

[14] William C. Y. Lee, "Smaller Cells for Greater Performance", *IEEE Communication. Magazine*, pp 19-23,November 1991.

[15] William C. Y. Lee, *Mobile Communication Design Fundamentals*,1993.

[16] David Macmillan, "Traffic Modelling and Analysis for Cellular Mobile Networks", *ITC-13*, pp 627-632, 1991.

[17] K. Maheshwari, "Performance Analysis of Microcellisation for Supporting Two Mobility Classes in Cellular Wireless Networks," Masters of Engineering Thesis, ECE Department, Indian Institute of Science, 1996.

[18] Sirin Tekinay and Bijan Jabbari, "Handover and Channel Assignment in Mobile Cellular Networks", *IEEE Communications Magazine*, pp 42-46, November 1991.

[19] Shashi B. Tripathi and Anurag Kumar, "Performance Analysis of Microcellisation with Channel Reservation for Supporting Two Mobility Classes in Cellular Wireless Networks," *IEEE Conference on Personal Wireless Communications*, December 1997, Mumbai, India, pp. 9-18.

[20] Shashi B. Tripathi, *Performance Analysis of Microcellisation with Channel Reservation in a Cellular Wireless Network*, Master of Engg. Thesis, ECE Department, Indian Institute of Science, Bangalore, June 1997.

[21] P.A.Whiting and D.M.McMillan, "Models for repacking in cellular radio", *7th UK teletraffic Symposium*, April 1990.

[22] K.L. Yeung and S.Nanda, "Optimal Mobile-Determined Micro-Macro Cell Selection", *IEEE International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC)*, pp. 294-299, 1995.