A Reversible Hierarchical Scheme for Microcellular Systems with Overlaying Macrocells

R. Beraldi, S. Marano, C. Mastroianni
D.E.I.S. Università della Calabria 87036 Rende (Cosenza) ITALY
tel: +39 984 494 702. e-mail: marano@ccuscl.unical.it

Abstract. Future cellular systems are expected to use multilayered, multisized cells to cover non-homogeneous populated areas. An example in literature is given by a 2 level hierarchical architecture in which an overlaying macrocell provides a group of overflow channels utilized when a microcell, which covers a densely populated area, is not able to accommodate a new call, or a handover from another microcell.

The macrocell has the higher hierarchical position, meaning that it can receive handover requests from microcells, lower in the hierarchy, as well as from other macrocells. On the contrary, a call served by the macrocell cannot handover to a microcell.

This paper proposes a Reversible Hierarchical Scheme characterized by the presence of handover attempts from macrocells to microcells. The scheme is conceived so that the microcells are given the majority of the traffic load as they are able to operate with very high capacity, while the macrocells, having lower channel utilization, can better carry out their support task.

An analytical study is carried out showing that the system performance can be improved, at the expense of relatively little increase of network control overhead, when compared with the classical, i.e. Non Reversible, Hierarchical Scheme.

1. Introduction

Conventional cellular radio systems employ relatively large cells, about 1-20 Km diameter, with antennas radiating relatively large powers (0.6 - 10W) from the top of tall buildings [1]. By the late 1990’s the demand for mobile radio services will overtake the capacity of second-generation systems technology (GSM, D-AMPS, etc.).

The most effective method of increasing the capacity of a cellular system is the use of microcells since, as cell sizes decrease, frequency reuse is possible within shorter distances. Microcells are usually less than 500 m in radius, thus requiring only low transmission power (i.e. less than 100 mW).

Several advantages of microcells are listed [2]:

- Significant increase in capacity, independent of RF standard.
- Non-high-speed pedestrian mobility simplifies signal processing.
- Low power allows miniaturization and more cost-effective radio design.
- Flexibility in RF design and teletraffic coverage.

However, microcells are not advantageous in service areas where user population density is sparse; moreover, small cell systems induce an increase in the number of cell boundary crossings (handovers) by mobile users. Micro-macrocell overlay structures, however, can overcome these difficulties (fig. 1). In future systems, there will be many types of cells whose size and shape are determined by the radiated power levels, the antenna location, and the user density in the area (typically the small size cells are used in the most populated areas). It is important to provide a balance between maximising the number of users per unit area (which favours small cells) and minimising the network control and handover rate (which favours large cells). The researchers have considered three important points in the design of the network [3]:

1. Determination of microcell size and position inside the macrocell; i. e. how to determine the traffic hot spots to be served by the microcells.
2. Frequency management and channel assignment (how many channels can be assigned to the microcells?).
3. Handover management.

The first design element, the location of the microcells, is to be determined based on the geographic configuration, traffic load, highway distribution, user density and some other related data.

Frequency management will be an important factor of successful microcell deployment; the goal is to find the most efficient way to allocate spectrum for different cell types. Bandwidth partitioning can be adopted. For example, microcells could be given the majority of the bandwidth as they are able to operate with very high capacity and will support the greatest variety of services. The overlaid macrocells will merely be used to serve the less populated areas, and to support handovers [4].
Finally, the handover management problem is to be considered. A handover occurs when a communicating Mobile Station (MS) leaves the cell in which it is currently being served, the source cell, and moves into another cell, the target cell. Handover calls between cells belonging to different hierarchical levels can remarkably improve the cellular system performances, and they must be carefully managed. In microcellular systems where many handovers are likely to occur during a call, it is essential to ensure that the probability of a call termination should be considerably smaller than the probability of blocking an attempted new call. This requirement is satisfied by the priority schemes, which reserve a certain number of channels to handovers ([1],[5]).

In [6] an analytical model for teletraffic performance analysis of hierarchically overlaid systems was developed. The approach is based on multidimensional birth-death processes, and on Markov analysis. In this model, a Fixed Channel Assignment (FCA) is considered, but macrocell channels can accommodate calls that cannot be served by the microcells. Performance characteristics that show carried traffic as well as blocking, handover failure and forced termination probabilities are derived from the analysis. The system operates according to a hierarchical scheme, so that a call served by the macrocell will not request handover to a cell that is lower in hierarchy, that is to a microcell: therefore, in this paper this scheme will be referred to as the Non Reversible Hierarchical (NRH) scheme.

However, this scheme could be inefficient in some situations: for example, suppose that many users originate new calls in the macrocell-only region (the region served only by the macrocell), and then move into a microcell: this can happen in the early morning, if the microcell covers an office area. The calls cannot request handover to the microcell, therefore the macrocell will be soon fully-loaded and several new calls will be blocked, even if the microcell has many idle channels. Moreover, with the hierarchical model, the microcells have lower channel utilization than the macrocell, but it should be the contrary: microcells should be given the majority of the traffic load as they are able to operate with very high capacity, while the macrocells should only provide overflow channels.

In this paper, a different way is proposed of managing the hierarchical microcell/macrocell architecture, and use a mathematical model similar to the one presented in [6] to derive the performance of the new system. In the system proposed, each time a communicating MS moves from the macrocell-only region to a microcell, a handover attempt is directed to the target microcell. If the microcell has any idle channel, then it will accommodate the handover attempt, otherwise no action will be taken, and the call will still be served by the macrocell: in any case, the call will not be forced to terminate. With this alternative network control, also referred to as Reversible Hierarchical (RH), the microcells tend to have higher channel utilization, and the whole system is more balanced in supporting the traffic load, as the performance indices will show. Additionally, a very important improvement is the considerable decrease of the blocking probability and the handover failure probability in the macrocell. The RH scheme performs even better when non-uniform channel allocation is utilized, that is when more channels are assigned to the macrocells than to the microcells.

The rest of the paper is organized as follows: in Sec. 2, the model of the Reversible Hierarchical scheme is given. Numerical results are presented in Sec. 3. Finally, conclusions are given in Sec. 4.

2. Model description

Early studies of cellular systems relied only on computer simulation, but some recent works ([6],[7]) developed an analytical model based on multidimensional birth-death processes. Since the development presented here proceeds along the lines of Reference [6], only a brief description of the model is presented, in order to concentrate on conceptual valuations.

The model description, based on the birth-death processes analysis, will be organized in four points:

1. Discussion of the assumptions and system parameters.
2. Identification of system states and driving processes.
3. Writing of the flow balance equations and determination of the equilibrium state probabilities.
4. Determination of the performance measures (blocking probability, handover failure probability, carried traffic).
2.1 System parameters and assumptions

The region under observation is covered by cells, called macrocells, each of which overlays several microcells. In a given macrocell, a part of the region is served only by the macrocell (this is the macrocell-only region), while the remaining area is covered by the macrocell and the microcells. The macrocell is denoted as cell 0, and the microcells as cell 1, 2, etc. The region is traversed by a large number of mobile users, pedestrians or vehicles, that have at most one call in progress. New calls and handover calls enter at both the microcell and macrocell level. The macrocell can accommodate the calls that cannot be served by the microcells: so the macrocell provides an overflow group of channels. A Fixed Channel Assignment (FCA) with cut-off priority scheme is used, so cell i is allocated a number of channels Ci, of which Chi are used only by handover calls (specific channels are not reserved, just the number). In this way a new call can use an idle channel in cell i, only if fewer than Ci - Chi channels are in use in that cell, while a handover call can use any idle channel: obviously, increasing Chi provides increasing priority for handovers at the expense of new call originations.

For convenience, it is assumed that the system is homogeneous in statistical equilibrium, i.e. any macro-area is statistically the same as any other macro-area. So the entire system will not be considered, but only a macro-area region, that contains a macrocell and the overlaid microcells.

It is assumed that the number of Mobile Stations in each cell is much larger than the number of channels allocated to that cell; each MS can support at most one call simultaneously. The amount of time that a communicating MS remains within the area covered by the Base Station (BS) of a cell is called dwell time. Dwell time depends on many factors, such as propagation conditions, the path a mobile platform follows, its velocity profile, the cell size, etc. Another parameter to be introduced is the unencumbered session duration of a call, which is defined as the amount of time that the call would remain in progress if it could continue to completion (without forced termination due to handover failure).

The following assumptions render the problem amenable to solution using the theory of multidimensional birth-death processes.

a) The new call arrival processes in any cell follow Poisson point processes. The new call arrival rate in the cell i, denoted by \( \lambda_i \), is the product of call demand rate from any MS and the number of MSs in the cell [7].

b) It is assumed that the handover arrival process from adjacent macro-areas follows a Poisson point process.

c) The dwell time of an MS in the cell i, denoted Di, is a random variable having a negative exponential probability density function, with mean 1/\( \mu_i \).

d) Also the unencumbered session duration of a call T is a random variable having a negative exponential pdf with a mean 1/\( \mu \).

The following analysis does not depend on any mobility model (that could be used in a computer simulation, see [8]), but the behaviour of the users and the geographical configuration of the macro-area are characterized by the teletraffic flow matrix, defined by:

\[
A = \begin{bmatrix}
\alpha_{00} & \alpha_{01} & \alpha_{02} & \ldots & \alpha_{04} \\
\alpha_{10} & \alpha_{11} & \alpha_{12} & \ldots & \alpha_{14} \\
\alpha_{20} & \alpha_{21} & \alpha_{22} & \ldots & \alpha_{24} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\alpha_{N0} & \alpha_{N1} & \alpha_{N2} & \ldots & \alpha_{NN}
\end{bmatrix}
\]

In the matrix A, the element \( \alpha_{ij}, i \neq j \), represents the probability that a handover departure from cell i is directed to cell j. For \( i = j \), \( \alpha_{ii} \) is defined as zero. The macrocell is numbered as "0", while the adjacent macro-areas are denoted as "d". It is important to note that for the Non Reversible Hierarchical scheme, the first row is not present, because there are no handover departures from the macrocell towards the microcells. For the Reversible Hierarchical scheme, on the other hand, the first row is necessary.

The teletraffic matrix A is assumed to be known and given, but it can be determined in a real macro-area under observation.

2.2 Identification of system states and driving processes

Since the system is homogeneous, only a macro-area will be considered and analyzed.

The macro-area under observation may be, in any instant, in any one of a finite number of states, depending on the number of channels assigned to communicating MS in the macrocell and in the N overlaid microcells. As in [6], the state is defined as a vector of non-negative integers (\( \nu_0, \nu_1, \ldots, \nu_N \)), where \( \nu_i \) (\( i=0,1,\ldots,N \)) is the number of calls served by the cell i.

A state is permissible if, \( \forall i = 0\ldots N \)

\[ 0 \leq \nu_i(s) \leq C_i \]

that is if in cell i the number of calls in progress does not exceed the number of channels Ci assigned to that cell.

1b.3.3
It is convenient to enumerate and order the states, in such a way that each (vector) state is assigned an integer index \( s \), ranging from 0 to \( S_{\text{max}} \), where \( S_{\text{max}}+1 \) is the number of possible states of the system:

\[
S_{\text{max}} + 1 = (C_0 + 1) \cdot (C_1 + 1) \cdots (C_x + 1)
\]

In this way, state \( s \) corresponds to the vector

\[
v(s) = v_0(s), v_1(s), \ldots, v_N(s)
\]

where \( v_i(s) \) is the number of calls served by the cell \( i \), when the system is in state \( s \). The total number of calls served in the macro-area is

\[
v(s) = \sum_{i=0}^{N} v_i(s)
\]

The system changes state at random instants, due to the following driving processes:

1) Generation of new calls in the macrocell-only region. A new call that originates in this region will be served by the macrocell if the number of idle channels in the macrocell is at least equal to \( C_{m0} \) (that is, the number of channels in use is fewer than \( C_0 - C_{m0} \)); otherwise the call will be blocked.

2) Generation of new calls in overlaid microcells. A new call that originates in the microcell \( i \) will be served by that microcell if the number of channels in use (\( V_i \)) is fewer than \( C_i - C_{m_i} \). Otherwise the call will be served by the macrocell if \( V_0 \) is fewer than \( C_0 - C_{m0} \); otherwise the call will be blocked.

3) Completion of calls in the macro-area under observation.

4) Handover arrivals to the macro-area from adjacent macro-areas. Such a handover call is always directed to the target macrocell, even if it impinges on a region covered by the microcell. The call will be served if there is any idle channel in the macrocell.

5) Handover arrivals to the macrocell-only region from overlaid microcells. The call is served if there is any idle channel in the macrocell, otherwise it will be blocked.

6) Handover departures from the macro-area under observation.

7) Handover arrivals to a microcell from adjacent microcells in the same macro-area. If the handover is from cell \( i \) to cell \( j \), it will be served by the microcell \( j \), if it has any idle channel, otherwise it will be served by the macrocell, if it has any idle channel, otherwise the handover attempt will fail.

8) Handover arrivals to a microcell from the macrocell-only region (this process is to be considered only in the RH scheme). This happens when an MS moves from the macrocell-only region to a region served by the macrocell and also by a microcell, say the microcell \( i \). The call will be served by the microcell \( i \), if it has any idle channel, otherwise it will continue to be served by the macrocell. This kind of handover attempt cannot fail.

### 2.3 Determination of the state probabilities

Since the state space is finite, the system can have an equilibrium distribution, that is a collection of positive numbers \( p(s), s = 0, \ldots, S_{\text{max}} \), summing to unity, that satisfy the flow balance equation for each state. So \( S_{\text{max}} + 1 \) linear equations for the unknown state probabilities are identified:

\[
\sum_{j=0}^{S_{\text{max}}} q(i,j) p(j) = 0, \quad i = 0,1,2,\ldots, S_{\text{max}}
\]

For \( i \neq j \), \( q(i,j) \) is the transition rate (or probability flow) from state \( j \) to state \( i \) (\( j \) is a predecessor of \( i \)), and \( q(i,i) \) is the total transition rate out of state \( i \). The convention is that the flow into a state is assumed positive. These equations are not linearly independent (the rank of the matrix is \( S_{\text{max}} \)), but a nonzero unique solution is determined using the condition:

\[
\sum_{s=0}^{S_{\text{max}}} p(s) = 1
\]

Now, the goal is to determine the transition rates \( q(i,j) \), i.e. the coefficients of the equilibrium equations (1). The approach followed to find these quantities is not presented here, because it is similar to that presented in [6]. Of course, appropriate modifications are necessary when considering the RH scheme.

Once all the coefficients needed in (1) have been calculated, the unknown state probabilities \( p(s) \) can be found solving the linear algebraic equations (1) and (2). Depending on the number of channels assigned to the cell, the number of macro-area states can be very large, so the algorithms and procedures used to solve the linear system may require much CPU time and memory space. However, the matrix related to the linear system is very sparse, i.e. only a relatively small number of its element are nonzero. Therefore if a numerical algorithm is chosen, that takes the maximum advantage of sparseness, a readily available workstation can take a reasonable time to solve the system.

### 2.4 Determination of performance measures

Three performance measures that characterize the cellular system are derived by the state probabilities: the
blocking probability, the handover failure probability and the average traffic carried.

First, it is useful to define an overall system parameter, the average transition rate or event rate, where an event may be a new call attempt, a handover request or a call termination. The event rate $R$ is given by

$$R = \sum_{s=0}^{s_{\text{max}}} |q(s,s)| \cdot p(s)$$

where $|q(s,s)|$ is the total transition rate into state $s$ (it is a positive quantity), and the equilibrium probability $p(s)$ is the fraction of time that the system spends in state $s$.

The blocking probability $P_{B}$ is defined as the average fraction of new call originations in the macro-area that cannot be served by the macrocell or the microcells. A new call originating in the macrocell-only region is blocked if $v_0(s) \geq C_0 - C_{10}$, that is the system is in one of the states belonging to the subset defined by

$$B_0 = \{ s : v_0(s) \geq C_0 - C_{10} \}$$

Therefore the blocking probability in the macrocell is

$$P_{B0} = \sum_{s \in B_0} p(s)$$

Analogously, a new call in the $i$th microcell ($i = 1, \ldots, N$) is blocked if $v_i(s) \geq C_i - C_{10}$ and $v_0(s) \geq C_0 - C_{10}$, that is the system is in one of the states of the subset:

$$B_i = \{ s : v_0(s) \geq C_0 - C_{10}, v_i(s) \geq C_i - C_{10} \}$$

The blocking probability in the $i$th microcell is given by

$$P_{Bi} = \sum_{s \in B_i} p(s) \quad i = 1, 2, \ldots, N$$

Since $B_i$ is a subset of $B_0$, the blocking probability in the macrocell is always greater than that in $i$th microcell. The reason is that while in the macrocell-only region a new call is given only one chance to access a channel, a new call originating in a microcell has two chances, because it can be accommodated in the microcell or in the macrocell.

The handover failure probability $P_{H}$ is defined as the average fraction of handover calls that fail to gain access to a channel in the target zone, causing the forced termination of the call. In [6] a simple way to calculate this performance measure was developed, based on the jump Markov chain probability distribution defined in [9].

The handover failure probability when a handover attempt impinges on the $i$th cell ($i = 0, 1, \ldots, N$) was found to be

$$P_{Hi} = \sum_{s \in H_i} p(s) \cdot \frac{\lambda_i}{R} \quad i = 0, \ldots, N$$

where $H_i$ is the subset of state in which a handover failure can occur in cell $i$; the parameter $\lambda_i$ is the rate of handover calls impinging on the $i$th cell, when the system is in state $s$, and it is given by

$$\lambda_i = \sum_{j=1}^{N} \mu_{ij} \cdot v_j(s) \cdot q_{ij} \quad i = 1, \ldots, N$$

The carried traffic in cell $i$, $AC_i$, $(i = 0, 1, \ldots, N)$ is the average number of channels in use in cell $i$, and it is given by

$$AC_i = \sum_{s=0}^{s_{\text{max}}} v_i(s) \cdot p(s) \quad i = 0, \ldots, N$$

where $v_i(s)$ is the number of occupied channels in cell $i$, when the system is in state $s$. The total carried traffic ($AC_{\text{total}}$) in the macro-area is

$$AC_{\text{total}} = \sum_{i=0}^{N} AC_i$$

3. Numerical Results

The performance measures of the Reversible Hierarchical scheme are obtained by using the model described for some possible hypothetical traffic conditions. To show the improvements of the RH scheme compared to the NRH scheme, it is opportune to analyze the same microcellular system that was considered in [6].

The sample macro-area consists of two similar microcells and an overlaying macrocell. Cell 0 is the macrocell, while the microcells are denoted as cell 1 and cell 2. System parameters and traffic conditions are the following:

- The mean unencumbered call duration is $T = 100$ sec.
- The mean dwell times in the macrocell and microcell are $T_{Do} = 225$ sec, $T_{Di} = T_{D2} = 150$ sec.

1b.3.5
The number of mobile stations (or users) in each cell is 550.

The number of channels allocated to the macro-area is 60. The channels are to be shared among the macrocell and the two microcells.

The teletraffic flow matrix is

$$A = \begin{bmatrix} 0.0 & 0.4 & 0.4 & 0.2 \\ 0.2 & 0.0 & 0.5 & 0.3 \\ 0.2 & 0.5 & 0.0 & 0.3 \end{bmatrix}$$

The first row of $A$ is only defined for the RH scheme: it is assumed that 40% of handover departures from the macrocell is towards each microcell (see the corresponding matrix elements $A_{01}$ and $A_{02}$), and 20% towards adjacent macro-areas (element $A_{0d}$).

In the first example, the call rate per user ranges from $7 \cdot 10^{-5}$ calls/sec (about one call each four hours) to $27 \cdot 10^{-5}$ calls/sec (about 1 call/hour). The offered traffic in each cell is calculated as the product of the call rate per user and the number of users in that cell. The channel allocation pattern is (20,20,20), that is 20 channels are assigned to the macrocell, and 20 to both microcells.

Fig. 2 shows the different load distribution in the macro-area for the NRH scheme and the RH scheme. By using the RH scheme, which allows handover calls from the macrocell to microcells, the traffic carried in the macrocell, $ACo$, is reduced, while the traffic served by each microcell (denoted as $AC_{1,2}$ because both microcells have the same statistical behaviour) is improved. The total traffic carried in the macro-area is nearly the same in both schemes, but for high new call rate, the RH scheme performs better than the NRH scheme, due to a more efficient channel utilization.

![Figure 2: Carried traffic in the macro-area. The channel allocation pattern is (20,20,20).](image)

Fig. 3 shows the originating call blocking probabilities in the macro-area. It is seen that $PBo$ is remarkably lower in the proposed RH scheme because of the lower channel utilization in the macrocell. The blocking probability of the microcells is also affected by the macrocell channel utilization, since a new originating call has two chances to find a channel: first, it tries in the microcell (which is more loaded than in the NRH scheme), then it tries in the macrocell (which has more idle channels than in the NRH scheme). The result is that the blocking probability in the microcells is decreased in the RH scheme for low offered traffic, while it is increased for a higher traffic. However the performance improvements in the macrocell, due to the RH scheme utilization is much more evident than the partial worsening in the microcells.

Similar behaviour is observed with regard to the handover failure probability (see Fig. 4). In the RH scheme, the performance curves of the macrocell and the microcells are quite similar, while in the NRH scheme there is a good performance of microcells but a poor one in the macrocell.

Fig. 5 shows the average event rate $R$ in the whole macro-area as a function of the call rate per user. The event rate in the RH scheme is increased (up to 27% for high offered traffic), resulting in a heavier network control needed to handle the events.

### 3.1 Non-uniform channel distribution

So far, the same number of channels has been allocated to the macrocell and both microcells. In the next examples, more channels are assigned to the macrocell and less to the microcells: channels allocation pattern is (32,14,14). Fig. 6 shows blocking probabilities in the macro-area for the NRH scheme and the RH scheme. The
new call rate per user was varied from $12 \cdot 10^{-5}$ to $27 \cdot 10^{-5}$ calls/sec, because negligible values of blocking probability were found for lower offered traffic. Again, a high improvement is observed in macrocell performance, when the RH scheme is used. With this channel pattern the blocking probability in the microcells is also decreased, even for high offered traffic, due to the increased number of overflow channels in the macrocell: thus, a new call whose access is denied to a microcell channel, has some more chances to find an idle channel in the macrocell. Similar trends can be observed from curves of handover failure probabilities, showed in Fig. 7.

As can be noted, with this non-uniform channel allocation, the RH scheme performs better than the NRH scheme in the microcells as well as in the macrocell.

It is certainly interesting to see how much the RH scheme performance is affected by the values of the teletraffic flow matrix elements. Fig. 8 shows the blocking probabilities in the macro-area versus the percentage of handover departures, from the macrocell, that flow towards the inner microcells (e.g. when the percentage is 60%, the matrix elements $a_{01}$ and $a_{02}$ are equal to 0.3, while $a_{03}=0.4$). The call rate per user is assumed to be $21 \cdot 10^{-5}$ calls/sec; the channel allocation pattern is $(32,14,14)$. It is seen that the blocking probabilities are decreased, in the macrocell as well as in the microcells, when the macrocell directs the most of the handover calls to the microcells. When the percentage is 0%, i.e. all the handover calls are directed to the adjacent macro-areas, the RH scheme coincides with the NRH scheme. In Fig. 9 the analogous curves of handover failure probabilities are displayed.

3.2 Effect of channel reservation

So far, no priority scheme has been adopted, that is no channels have been reserved to handover calls. In [6] it
was shown that the use of channel reservation improves the system performance with regard to handover calls management, but the blocking probability is increased because a lower number of channels can be used by originating calls. The number of channels reserved to handovers is an important design parameter, and it must be carefully chosen according to specific network requirements.

4. Conclusions

A reversible hierarchical architecture is presented and analyzed in the paper. In this scheme the overlaying macrocell provides a group of overflow channels, and furthermore it allows handover attempts from the macrocell to the microcells to take place, in order to partially unload the macrocell and to obtain a higher channel utilization in the microcells.

Due to the higher probability of finding free channels in the macrocell, both blocking probability and handover failure probability are decreased but a slightly higher control overhead is revealed.

In this way the macrocell channels are utilized for a shorter amount of time, therefore they can be used more effectively as overflow channels by a larger number of users.

The scheme could be generalized, that is the cellular system can have more levels in the hierarchy (e.g. the most populated areas could be covered by macrocells, microcells, and picocells too).

REFERENCES


