TRAFFIC MODELLING AND ANALYSIS FOR CELLULAR MOBILE NETWORKS

David McMILLAN

Telecom Australia Research Laboratories, P.O. Box 249
Melbourne, 3168, Australia

This paper proposes a model for cellular mobile traffic which takes into account the mobility of users and hence their handover requirements. The traffic model is simple enough to allow performance parameters to be estimated analytically and for it to be constructed from data available in the current mobile switching centres. Under further statistical assumptions about the traffic model, approximations for network control strategies in terms of fresh call blocking and handover blocking can be constructed. An example is given of a control strategy with fixed channel assignment, directed retry for fresh call attempts and channel reservation for handoff attempts. The channel reservation strategy gives priority to handoff attempts over fresh call attempts. The main contribution of this paper is a methodology for the approximate analysis of control strategies in cellular mobile networks.

1 Introduction

The distinguishing features of cellular mobile networks which set them apart from conventional circuit-switched networks are channel re-use and the handover facility. The feature of channel re-use allows for cellular networks to increase capacity by reducing cell sizes. The handover facility allows for customers to roam while they have a call in progress. Both of these features create problems for network designers and managers. Channel re-use requires careful management of both co-channel and adjacent-channel interference. The roaming of customers has two major effects: firstly, it causes a changing interference situation which must be allowed for in the channel assignment scheme to prevent call dropouts and, secondly, it creates the need for handovers which should succeed with a high probability.

A network control strategy can be considered to have three parts: channel assignment; call acceptance and the handover strategy. The channel assignment strategy determines which channels are in use in each cell. The call acceptance strategy determines to which cell a fresh call should be connected and, if necessary, re-arranges calls in progress. The handover strategy determines to which cell a handover attempt should connect and the way in which the handover is performed. The three strategies are not independent since, for example, a handover attempt may require the re-arrangement of calls in progress or it may require a channel re-allocation. Channel assignment, call acceptance and handover strategies are classified more extensively in reference [8].

There is no attempt to model the changing interference situation due to call arrivals and customer mobility. Instead, it is assumed that interference is accounted for implicitly in the channel assignment. It is expected that there may be unexpected dropouts and/or a waste of capacity due to this (but this is the case for current and near-future systems anyway). A function of network control, not listed previously, but which further improves network performance is that of power adaptation [1]. Power adaptation strategies control the power of both the mobile and base station transmitters in an effort to reduce interference levels. There is no attempt to model these strategies, but it is expected that their use may be partly accounted for in the channel allocation schemes.

In this paper, a traffic model is proposed which is simple enough to be used in the analysis of the above-mentioned strategies. The model takes into account the mobility of users and hence handover requirements and allows for control strategies to be evaluated in terms of performance parameters such as fresh call blocking, handover blocking and handover delay. The traffic model is simple enough to be easily constructed and allows for an analytical estimate of performance parameters thus giving insight into control strategies. Information that is required in the traffic model is: fresh call arrival rates to the cells; handover rates to neighbouring cells (per user); and call holding times.

The handover rates specified above are averages and are preferred on a per user basis so that they are independent of the arrival rates. They could alternatively be specified as handover probabilities, a handover probability being the probability that a user will request a
handover to a particular neighbour. It should be noted that handover rates are dependent on the form of the cellular network through the size and shape of each cell and in particular vehicular traffic patterns in that cell. The traffic model allows for the cells to be of any shape and to have any number of neighbours. The required information is already available in the mobile switching centres for current networks and could also be estimated if there were no network data.

With the above-mentioned traffic model, the performance of network control strategies can be estimated by modelling the cellular network as a network of queues; as described in Section 2. In general, it is difficult to solve for the equilibrium distribution describing the state of the network and approximations are usually required. In Section 3, an approximation technique is described for the simple case: a network with Fixed Channel Assignment (FCA) and equal blocking (loss) for fresh call attempts and handover attempts. In Section 4, the approximation technique developed in Section 3 is extended in order to model channel reservation, a handover strategy which gives priority to handover attempts over fresh call attempts. In Section 5, the technique is extended to include directed retry, a call acceptance strategy which exploits cell overlap areas. In Section 6, numerical examples are presented for a model network of 49 cells and an offered traffic profile with a peak in the centre cell.

In [3], a fixed point model is developed under similar assumptions used in this paper and which gives the same result for FCA with channel reservation. This paper attempts a more formal argument for the procedure and includes analysis of the directed retry strategy.

2 Cellular network as an open queueing network

At first, we will think of a cellular network as an open queueing network. For each cell (queue) there are fresh call arrivals and handover arrivals from neighbouring cells. Handovers are modelled as transitions between queues: after spending a certain amount of time in a cell (queue), a customer may either terminate the call (leave the system) with a certain probability or attempt a handover to another cell (a transition) with a certain probability. The time spent in a cell before either call termination or handover is assumed to have a negative exponential distribution.

This open queueing network will have the following description: a total of \( J \) cells in the network; users arrive to queue \( j \) as a Poisson process of rate \( \nu_j \); a user will wish to move from cell \( j \) to cell \( k \) at rate \( \lambda_{jk} \) (per unit time); a user will wish to leave the network at rate \( \mu \) (per unit time).

The total rate of departure from cell \( j \) will be given by:

\[ \lambda_j = \mu + \sum_k \lambda_{jk}. \]  

A user will require service from cell \( j \) for a time which is negative–exponentially distributed with mean \( 1/\lambda_j \). This is the channel holding time. After receiving this service the user will move to cell \( k \) with probability \( \lambda_{jk}/\lambda_j \). The call holding time (as distinct from the channel holding time), for a call which does not suffer internal blocking in the network, has a negative exponential distribution with rate \( \mu \). An assumption which is inherent in the above description is that users move at random through the service area according to a set of probabilities: i.e. it is a one-step (Markov) probability model.

In this section the rate and traffic equations will be stated since they lead naturally to the approximation methods used in later sections. A discussion of the existence of exact analytic solutions for open queueing networks in certain special cases can be found, for example, in Kelly [4].

The rate equations are found as follows: suppose that the network is in equilibrium and let \( \alpha_j \) be the average number of jobs arriving into (and leaving) queue \( j \) per unit time. Then we have the following rate equation due to conservation of job flow:

\[ \alpha_j = \nu_j + \sum_k \alpha_k \frac{\lambda_{kj}}{\lambda_k}. \]  

The average number of users in cell \( j \), \( \rho_j = \alpha_j/\lambda_j \) by Little's law [4]. Thus, we have the following traffic equation:

\[ \rho_j \lambda_j = \nu_j + \sum_k \rho_k \lambda_{kj}. \]  

Both sets of equations are known to have a unique solution [4].

Using the traffic model described above, we want to evaluate the performance of the network under various control strategies. Channel assignment strategies correspond to capacity constraints on the queues (cells). Handover strategies will correspond to restrictions on movements between cells. When there is no blocking in the network, analytic expressions can often be found for the equilibrium probabilities. When there is blocking, however, analytic expressions are elusive and approximations need to be made.

3 Fixed point approximations for simple blocking networks

For our purposes we will define a simple blocking network as one in which all users attempting to find a channel in a particular cell, whether they are fresh attempts or handover attempts, suffer the same level of blocking. If a handover attempt is blocked then the call is lost. In addition, the number of channels in each cell is fixed. Development of the fixed-point approximation for calculating blocking probabilities for
this simple network follows naturally from the queueing network theory in the previous section and also leads to approximation methods for improved channel assignment and handover strategies.

The fixed-point approximation is based upon three assumptions: that fresh call arrivals are mutually independent; that cell blocking probabilities are mutually independent; and that handovers arrive to each cell as Poisson processes. None of these premises will be true in general but they allow for the blocking values to be calculated from simple queueing models. This justification is the same as that used for the Erlang fixed point approximation which estimates blocking values for circuit switched networks with fixed routing [5].

Let \( \nu_j, \lambda_j \) and \( \lambda_j \) be as defined in the previous section and suppose that: \( \alpha_j \) is the steady state rate in and out of cell \( j \) (as before); and \( B_j \) is the blocking probability for cell \( j \).

Then if the network is in equilibrium, conservation of flow gives that the total offered rate to cell \( j \), \( \theta_j \), will be the sum of the fresh arrival rate and the total handover rate:

\[
\theta_j = \nu_j + \sum_k \alpha_k \frac{\lambda_{kj}}{\lambda_k}.
\]

Under the assumption that offered traffics are Poisson and cell blockings are independent,

\[
\alpha_j = (1 - B_j)\theta_j.
\]

If we assume that the offered calls to cell \( j \) are from a Poisson process of rate \( \theta_j \), and that we are using fixed channel assignment, the blocking probability for cell \( j \) is given by:

\[
B_j = E(\theta_j, C_j)
\]

where \( C_j \) is the number of channels in cell \( j \); and \( E(\cdot, \cdot) \) is Erlang’s loss function [6].

Note that it has been assumed that both fresh and handover traffic have the same channel holding time after acceptance in a particular cell. Further, the transition rates are considered to be the same for both types of traffic.

In terms of the offered traffics, \( \rho_j = \theta_j/\lambda_j \), equations (4),(5) give:

\[
\rho_j \lambda_j = \nu_j + \sum_k \rho_k \lambda_{kj} (1 - B_k),
\]

These equations have a unique solution for \( B_j < 1 \) since they can be recast in the form of equation (2) by modifying the transition rates. We can write these equations as a linear system,

\[
(\Lambda_r - R^T \Lambda_b)\rho = \nu,
\]

so that

\[
\rho = A^{-1}\nu,
\]

where \( \Lambda_r = \text{diag}(\lambda_j) \), \( \Lambda_b = \text{diag}(1 - B_j) \), \( R = (\lambda_j) \), \( \nu = (\nu_j) \), \( \rho = (\rho_j) \), \( A = \Lambda_r - R^T \Lambda_b \)

The above equations define a fixed point for the blocking values:

\[
B_j = E(\rho_j(B), C_j) \quad j = 1, \ldots, J.
\]

The equations define a mapping from \([0,1]^J\) to \([0,1]^J\) and so, by the Brouwer fixed point theorem, a fixed point exists. In practice, the blocking values can be found iteratively by choosing an initial vector of blocking values, then using fixed point iteration with the above equations.

Consider the linear system in equation (8). By expanding the inverse, we can see the effects of multiple handovers on the offered traffic to each cell.

Define \( X = R^T \Lambda_b \Lambda_r^{-1} \), and since \( A = (I - X)\Lambda_r \), then

\[
\rho = A^{-1}\nu = \Lambda_r^{-1} (I - X)^{-1} \nu
\]

\[
= \Lambda_r^{-1} \left( I + \sum_{i=1}^{\infty} X^i \right) \nu.
\]

Suppose we truncate the infinite sum contained within the brackets in the equation above. If we truncate the sum to just the identity matrix then we get the equations:

\[
\rho_j \lambda_j = \nu_j,
\]

in which case the offered traffic is just the fresh traffic and there has been no accounting for handovers. If we truncate the sum to, say, \((I + X + X^2)\), then we get the equations:

\[
\rho_j \lambda_j = \nu_j + \sum_k \nu_k \frac{\lambda_{kj}}{\lambda_k} (1 - B_k) + \sum_i \sum_k \nu_i \frac{\lambda_{ij}}{\lambda_i} \lambda_{kj} (1 - B_k) (1 - B_k),
\]

in which case we have accounted for fresh traffic plus handover traffic due to up to two handovers.

From the truncations we see that we could define a more sophisticated traffic model. For instance, if we wished to limit the number of handovers to two, then we could use equation (12) to find the offered traffics. Alternatively, we could define a model with more general customer routes than the one-step probability model used, but this is probably unnecessary and it would be a less manageable model.

### 4 Handover priority mechanisms

The fixed-point approximation developed in the last section does not distinguish between fresh call attempts and handover attempts in a given cell. Since handover attempts compete with fresh call attempts for channels in each cell, the handover attempts need to be given some sort of priority. Handover strategies are designed primarily to give very low handover blocking.
Some delay is allowable in the handover procedure, although, as cell sizes become smaller with digital technology, handover attempt rates will increase and there will be less tolerance for delay in the handover procedure. Handover strategies will need to be carefully designed in order to solve these problems and to provide for low handover blocking.

There are many ways in which to give priority to handover attempts over fresh call attempts. With channel reservation, a number of channels in each cell are reserved for access by handed-over calls only. This is a simple scheme to implement and is very effective. The only problem is that it can be wasteful if handover traffic is low for any length of time. An alternative strategy might be to place handover requests for a particular cell in a queue. Under this strategy, a new call would be offered a free channel only if there are no outstanding handover requests. A queueing strategy has the potential to provide greater system capacity than a reservation strategy but the delay must be considered. A strategy combining channel reservation and queueing could also be employed. The channel reservation strategy is modelled in this paper; the queueing option can be modelled in similar fashion.

### 4.1 Channel reservation priority

To derive a fixed-point formulation for modelling channel reservation, a similar assumption is made to that made previously, i.e., call blocking and handover blocking probabilities are independent from cell to cell. Blocking values are calculated using a simple queueing model for each cell which assumes that the fresh and handover attempts are generated by independent Poisson processes.

Suppose that:

- $B_f^j$ is the fresh call blocking probability for cell $j$;
- $B_h^j$ is the handover call blocking probability for cell $j$;
- $\alpha_j$ is the steady-state rate of calls in (and out) of cell $j$;
- $\rho_{fj}$ is the offered fresh traffic for cell $j$;
- $\rho_{hj}$ is the offered handover traffic for cell $j$;
- $\gamma_j$ is the total carried traffic for cell $j$;
- $C_j$ is the number of channels for cell $j$;
- $t_j$ is the number of channels reserved in cell $j$ for access by handover attempts only (the channel reservation parameter).

A fresh call is accepted if there are less than $(C_j - t_j)$ circuits occupied; otherwise it is blocked. A handover attempt is accepted if there are any circuits unoccupied; otherwise it is blocked. It is assumed that channel holding times have a negative exponential distribution. Under this model,

$$B_f^j = E^f(\rho_{fj}, \rho_{hj}, C_j, t_j)$$
$$B_h^j = E^h(\rho_{fj}, \rho_{hj}, C_j, t_j)$$

where, e.g. [5],

$$E^f(\rho_f, \rho_h, c, t) = G(\rho_f, \rho_h, c, t) \frac{(\rho_f + \rho_h)^{c-t} \rho_h^t}{c!}$$
$$E^h(\rho_f, \rho_h, c, t) = G(\rho_f, \rho_h, c, t)(\rho_f + \rho_h)^{c-t} \sum_{n=c-t}^{c} \frac{\rho_h^n}{n!}$$

$$G^{-1}(\rho_f, \rho_h, c, t) = \sum_{n=0}^{c-t-1} \frac{(\rho_f + \rho_h)^n}{n!} + (\rho_f + \rho_h)^{c-t} \sum_{n=c-t}^{c} \frac{\rho_h^n}{n!}.$$

If we assume independent blocking, then in equilibrium:

$$\alpha_j = (1 - B_f^j)\nu_j + (1 - B_h^j)\sum_k \alpha_k \frac{\lambda_{kj}}{\lambda_k}. \quad (13)$$

Since the channel holding time for both traffics is the same, we can write by Little's law that:

$$\gamma_j = \frac{\alpha_j}{\lambda_j}$$
$$\rho_{fj} = \frac{\nu_j}{\lambda_j}$$
$$\rho_{hj} = \frac{1}{\lambda_j} \sum_k \gamma_k \lambda_{kj}.$$  

The carried traffic equations can be solved as a linear system:

$$\gamma = A^{-1}A_{yk}\nu$$  

(14)

where $\gamma = (\gamma_j)$, $A_{yk} = \text{diag}(1 - B_f^j)$, $A_{yk} = \text{diag}(1 - B_h^j)$, $A = A_y - RTA_{hk}$, and $A_{yk}, R$ are as defined in the previous section. The blocking probabilities can be solved for iteratively as in the last section. Also, as in the last section, at least one fixed point is known to exist.

### 4.2 Calculation of forced termination probabilities

In Section 3, call blocking and handover blocking probabilities were calculated. Another parameter of interest is the probability of forced termination (drop-out) due to handover failure, given that any call may require several handovers. A good way in which to measure this is to calculate the forced termination probability, $B_f^j$, for a call originating in cell $j$. This is done by considering all possible routes through the network that the user may take.
Suppose that a call starts in cell \( j \). With probability \( \frac{\lambda_i}{\lambda_j} B_i^h \) the call will attempt to handover to cell \( i \) and fail, while with probability \( \frac{\lambda_i}{\lambda_j} (1 - B_i^h) \) it will succeed in handing over to cell \( i \).

Let \( p_i^k \) be the probability of failure for a call starting in cell \( j \) on the \( k \)th handover. Then

\[
p_i^1 = \sum \frac{\lambda_i}{\lambda_j} B_i^h
\]

\[
p_i^{k+1} = \sum \frac{\lambda_i}{\lambda_j} (1 - B_i^h) p_i^k.
\]

In matrix form, with \( B^h = (B_i^h) \), \( p^h = (p_i^h) \), \( B^d = (B_j^d) \), and \( X = \Lambda_0^{-1} R \Lambda_{hh} \) we have:

\[
p^1 = \Lambda_0^{-1} R B^h
\]

\[
p^{k+1} = X^k p^1.
\]

The dropout probabilities are then:

\[
B^d = \sum_{k=1}^{\infty} p^k
\]

\[
= \left( I + \sum_{k=1}^{\infty} X^k \right) p^1
\]

\[
= (I - X)^{-1} p^1.
\]

Finally, the probability of a call terminating successfully, given that it originated in cell \( j \), is:

\[
(1 - B_j^d)(1 - B_j^f).
\]

5 Directed retry

In order to ensure complete coverage of a cellular service area, there will be a large proportion of this area in which a mobile will be able reliably to connect with any one of two or more base stations. Studies have shown that this overlap proportion is in the range of 20–40% [7], depending on the cell layout. The directed retry strategy exploits the overlap area by allowing mobiles in an overlap area at call setup to connect with any one of the base stations involved in the overlap area. It is classed as an alternative routing strategy and provides a way of improving channel utilisation and hence system capacity. Usually, a mobile will first attempt to connect with the base station which provides the strongest signal. There are variations on this method: for example, the mobile may first attempt to connect with the base station which has the greatest spare capacity.

For the analysis, it is assumed that: call attempts arrive to cell \( i \) as a Poisson process with rate \( \nu_i \), these processes being independent. It is also assumed a fraction \( f_{ij} \) of call attempts to cell \( i \) are able to connect with cell \( j \). If they are blocked at cell \( i \), they then attempt to connect with cell \( j \).

To derive a fixed-point formulation we assume that the call blocking probabilities are independent from cell to cell. Given that the fresh blocking probability to cell \( i \) is \( B_i^f \) then the offered fresh call rate to cell \( j \) in equilibrium will be the sum of the first attempt rate and the directed retry rate:

\[
\theta_i^f = \nu_i + \sum_{j} \nu_j f_{ij} B_j^f.
\]

If \( \theta_i^f = (\theta_i^f) \) is the vector of offered fresh call rates, then

\[
\theta_i^f = F \nu
\]

where

\[
F_{ij} = f_{ij} B_j^f, \quad i \neq j \]

\[
F_{ii} = 1
\]

The fixed point derivation for the case of mobility is found by replacing \( \nu \) in the previous sections, wherever it appears, with \( \theta_i^f = F \nu \). The average fresh call blocking for a cell needs to take account of users in overlap areas and users who are not in overlap areas. Assuming independent blocking, this will be given by,

\[
B_i^{f*} = \left( 1 - \sum_{j} f_{ij} \right) B_i^f + \sum_{j} f_{ij} B_j^d B_j^f
\]

The handover strategy does not make use of the cell overlap area, unlike the call acceptance policy. It also assumes that a user ‘belongs’ to the cell to which it is connected, not the cell in which it originated.

6 Example

In this example, the directed retry strategy with channel reservation is evaluated for a model network of 49 cells. The offered traffics are shown in Figure 1 with a peak in the centre cell of 50 Erlangs. The handover rates have a peak of 0.2 for the centre cell and fall away in value towards the edges of the network (in the same way as the offered traffics). The rate of call departures, \( \mu_i \), is set to unity. There are 50 channels in every cell.

In Figure 2, both fresh and handoff blocking figures are shown for the following cases: FCA with no priority; FCA with channel reservation of 2 channels; FCA with directed retry for an overlap of 5% and with channel reservation of 2 channels. The distance from the centre is measured in numbers of cells, i.e., the centre cell is at a distance of 0 and the boundary cells are at a distance of 3. Note that for the first case the fresh and handover blocking values are equal. In the third case, the results are compared with simulation (see confidence intervals) and are shown to be accurate. With the introduction of channel reservation, the handover blocking probabilities have dropped to acceptable levels, causing a necessary increase in fresh blocking probabilities. The use of directed retry significantly improves the fresh blocking values.
7 Conclusion

In this paper, a traffic model has been introduced which takes into account the mobility of users and hence handover requirements. This model is simple enough to allow for analysis of network control strategies, thus giving insight into the behaviour of the network. It is particularly important to be able to model realistic handover strategies.

More complex control strategies will be more difficult to model than those discussed in this paper but should still yield to an approximate analysis perhaps under further simplifying assumptions. The proposed traffic model and analysis of simple strategies in this paper provides a suitable foundation for analysis of these more complex strategies.

Acknowledgements

The permission of the Executive General Manager, Research, of Telecom Australia to publish this paper is hereby acknowledged. The work described in this paper forms a part of the author's PhD studies undertaken at Melbourne University. The author also wishes to thank David Everitt for his helpful comments.

References


