

CHAPTER 1**Handoff in Wireless Mobile Networks**

QING-AN ZENG and DHARMA P. AGRAWAL

*Department of Electrical Engineering and Computer Science,
University of Cincinnati***1.1 INTRODUCTION**

Mobility is the most important feature of a wireless cellular communication system. Usually, continuous service is achieved by supporting handoff (or handover) from one cell to another. Handoff is the process of changing the channel (frequency, time slot, spreading code, or combination of them) associated with the current connection while a call is in progress. It is often initiated either by crossing a cell boundary or by a deterioration in quality of the signal in the current channel. Handoff is divided into two broad categories—hard and soft handoffs. They are also characterized by “break before make” and “make before break.” In hard handoffs, current resources are released before new resources are used; in soft handoffs, both existing and new resources are used during the handoff process. Poorly designed handoff schemes tend to generate very heavy signaling traffic and, thereby, a dramatic decrease in quality of service (QoS). (In this chapter, a handoff is assumed to occur only at the cell boundary.) The reason why handoffs are critical in cellular communication systems is that neighboring cells are always using a disjoint subset of frequency bands, so negotiations must take place between the mobile station (MS), the current serving base station (BS), and the next potential BS. Other related issues, such as decision making and priority strategies during overloading, might influence the overall performance.

In the next section, we introduce different types of possible handoffs. In Section 1.3, we describe different handoff initiation processes. The types of handoff decisions are briefly described in Section 1.4 and some selected representative handoff schemes are presented in Section 1.5. Finally, Section 1.6 summarizes the chapter.

1.2 TYPES OF HANDOFFS

Handoffs are broadly classified into two categories—hard and soft handoffs. Usually, the hard handoff can be further divided into two different types—intra- and intercell handoffs. The soft handoff can also be divided into two different types—multiway soft handoffs and softer handoffs. In this chapter, we focus primarily on the hard handoff.

2 HANDOFF IN WIRELESS MOBILE NETWORKS

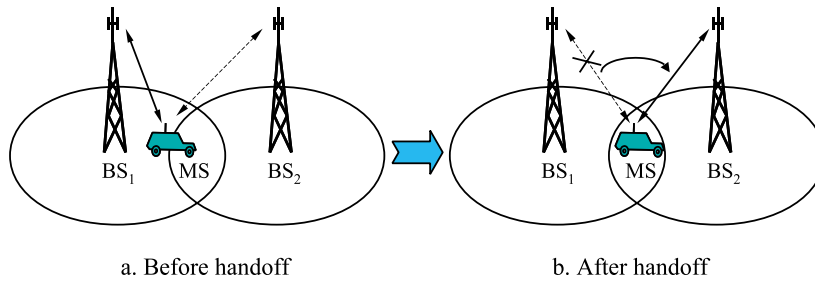


Figure 1.1 Hard handoff between the MS and BSs.

A hard handoff is essentially a “break before make” connection. Under the control of the MSC, the BS hands off the MS’s call to another cell and then drops the call. In a hard handoff, the link to the prior BS is terminated before or as the user is transferred to the new cell’s BS; the MS is linked to no more than one BS at any given time. Hard handoff is primarily used in FDMA (frequency division multiple access) and TDMA (time division multiple access), where different frequency ranges are used in adjacent channels in order to minimize channel interference. So when the MS moves from one BS to another BS, it becomes impossible for it to communicate with both BSs (since different frequencies are used). Figure 1.1 illustrates hard handoff between the MS and the BSs.

1.3 HANDOFF INITIATION

A hard handoff occurs when the old connection is broken before a new connection is activated. The performance evaluation of a hard handoff is based on various initiation criteria [1, 3, 13]. It is assumed that the signal is averaged over time, so that rapid fluctuations due to the multipath nature of the radio environment can be eliminated. Numerous studies have been done to determine the shape as well as the length of the averaging window and the older measurements may be unreliable. Figure 1.2 shows a MS moving from one BS (BS1) to another (BS2). The mean signal strength of BS1 decreases as the MS moves away from it. Similarly, the mean signal strength of BS2 increases as the MS approaches it. This figure is used to explain various approaches described in the following subsection.

1.3.1 Relative Signal Strength

This method selects the strongest received BS at all times. The decision is based on a mean measurement of the received signal. In Figure 1.2, the handoff would occur at position A. This method is observed to provoke too many unnecessary handoffs, even when the signal of the current BS is still at an acceptable level.

1.3.2 Relative Signal Strength with Threshold

This method allows a MS to hand off only if the current signal is sufficiently weak (less than threshold) and the other is the stronger of the two. The effect of the threshold depends

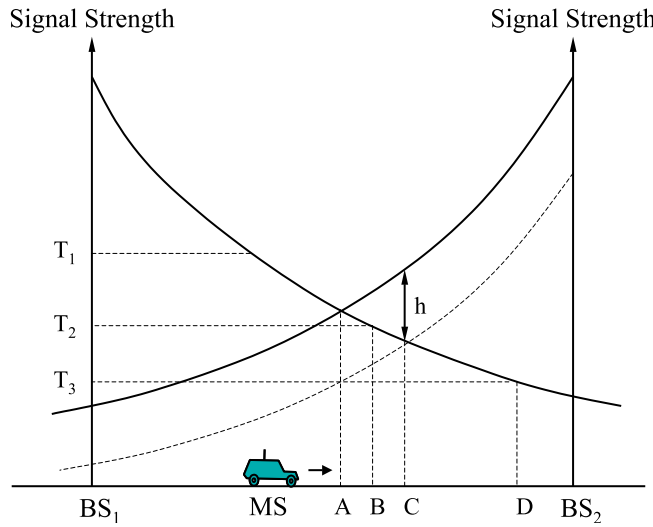


Figure 1.2 Signal strength and hysteresis between two adjacent BSs for potential handoff.

on its relative value as compared to the signal strengths of the two BSs at the point at which they are equal. If the threshold is higher than this value, say T_1 in Figure 1.2, this scheme performs exactly like the relative signal strength scheme, so the handoff occurs at position A. If the threshold is lower than this value, say T_2 in Figure 1.2, the MS would delay handoff until the current signal level crosses the threshold at position B. In the case of T_3 , the delay may be so long that the MS drifts too far into the new cell. This reduces the quality of the communication link from BS₁ and may result in a dropped call. In addition, this results in additional interference to cochannel users. Thus, this scheme may create overlapping cell coverage areas. A threshold is not used alone in actual practice because its effectiveness depends on prior knowledge of the crossover signal strength between the current and candidate BSs.

1.3.3 Relative Signal Strength with Hysteresis

This scheme allows a user to hand off only if the new BS is sufficiently stronger (by a hysteresis margin, h in Figure 1.2) than the current one. In this case, the handoff would occur at point C. This technique prevents the so-called ping-pong effect, the repeated handoff between two BSs caused by rapid fluctuations in the received signal strengths from both BSs. The first handoff, however, may be unnecessary if the serving BS is sufficiently strong.

1.3.4 Relative Signal Strength with Hysteresis and Threshold

This scheme hands a MS over to a new BS only if the current signal level drops below a threshold and the target BS is stronger than the current one by a given hysteresis margin. In Figure 1.2, the handoff would occur at point D if the threshold is T_3 .

4 HANDOFF IN WIRELESS MOBILE NETWORKS

1.3.5 Prediction Techniques

Prediction techniques base the handoff decision on the expected future value of the received signal strength. A technique has been proposed and simulated to indicate better results, in terms of reduction in the number of unnecessary handoffs, than the relative signal strength, both without and with hysteresis, and threshold methods.

1.4 HANDOFF DECISION

There are numerous methods for performing handoff, at least as many as the kinds of state information that have been defined for MSs, as well as the kinds of network entities that maintain the state information [4]. The decision-making process of handoff may be centralized or decentralized (i.e., the handoff decision may be made at the MS or network). From the decision process point of view, one can find at least three different kinds of handoff decisions.

1.4.1 Network-Controlled Handoff

In a network-controlled handoff protocol, the network makes a handoff decision based on the measurements of the MSs at a number of BSs. In general, the handoff process (including data transmission, channel switching, and network switching) takes 100–200 ms. Information about the signal quality for all users is available at a single point in the network that facilitates appropriate resource allocation. Network-controlled handoff is used in first-generation analog systems such as AMPS (advanced mobile phone system), TACS (total access communication system), and NMT (advanced mobile phone system).

1.4.2 Mobile-Assisted Handoff

In a mobile-assisted handoff process, the MS makes measurements and the network makes the decision. In the circuit-switched GSM (global system mobile), the BS controller (BSC) is in charge of the radio interface management. This mainly means allocation and release of radio channels and handoff management. The handoff time between handoff decision and execution in such a circuit-switched GSM is approximately 1 second.

1.4.3 Mobile-Controlled Handoff

In mobile-controlled handoff, each MS is completely in control of the handoff process. This type of handoff has a short reaction time (on the order of 0.1 second). MS measures the signal strengths from surrounding BSs and interference levels on all channels. A handoff can be initiated if the signal strength of the serving BS is lower than that of another BS by a certain threshold.

1.5 HANDOFF SCHEMES

In urban mobile cellular radio systems, especially when the cell size becomes relatively small, the handoff procedure has a significant impact on system performance. Blocking

probability of originating calls and the forced termination probability of ongoing calls are the primary criteria for indicating performance. In this section, we describe several existing traffic models and handoff schemes.

1.5.1 Traffic Model

For a mobile cellular radio system, it is important to establish a traffic model before analyzing the performance of the system. Several traffic models have been established based on different assumptions about user mobility. In the following subsection, we briefly introduce these traffic models.

1.5.1.1 Hong and Rappaport's Traffic Model (Two-Dimensional)

Hong and Rappaport propose a traffic model for a hexagonal cell (approximated by a circle) [5]. They assume that the vehicles are spread evenly over the service area; thus, the location of a vehicle when a call is initiated by the user is uniformly distributed in the cell. They also assume that a vehicle initiating a call moves from the current location in any direction with equal probability and that this direction does not change while the vehicle remains in the cell.

From these assumptions they showed that the arrival rate of handoff calls is

$$\lambda_H = \frac{P_h(1 - B_O)}{1 - P_{hh}(1 - P'_f)} \lambda_O \quad (1.1)$$

where

P_h = the probability that a new call that is not blocked would require at least one handoff

P_{hh} = the probability that a call that has already been handed off successfully would require another handoff

B_O = the blocking probability of originating calls

P'_f = the probability of handoff failure

λ_O = the arrival rate of originating calls in a cell

The probability density function (pdf) of channel holding time T in a cell is derived as

$$f_T(t) = \mu_C e^{-\mu_C t} + \frac{e^{-\mu_C t}}{1 + \gamma_C} [f_{T_n}(t) + \gamma_C f_{T_h}(t)] - \frac{\mu_C e^{-\mu_C t}}{1 + \gamma_C} [F_{T_n}(t) + \gamma_C F_{T_h}(t)] \quad (1.2)$$

where

$f_{T_n}(t)$ = the pdf of the random variable T_n as the dwell time in the cell for an originated call

$f_{T_h}(t)$ = the pdf of the random variable T_h as the dwell time in the cell for a handed-off call

$F_{T_n}(t)$ = the cumulative distribution function (cdf) of the time T_n

$F_{T_h}(t)$ = the cdf of the time T_h

6 HANDOFF IN WIRELESS MOBILE NETWORKS

$1/\mu_C$ = the average call duration
 $\gamma_C = P_h(1 - B_O) / [1 - P_{hh}(1 - P_f)]$

1.5.1.2 El-Dolil et al.'s Traffic Model (One-Dimensional)

An extension of Hong and Rappaport's traffic model to the case of highway microcellular radio network has been done by El-Dolil et al. [6]. The highway is segmented into microcells with small BSs radiating cigar-shaped mobile radio signals along the highway. With these assumptions, they showed that the arrival rate of handoff calls is

$$\lambda_H = (R_{cj} - R_{sh})P_{hi} + R_{sh}P_{hh} \quad (1.3)$$

where

P_{hi} = the probability that a MS needs a handoff in cell i
 R_{cj} = the average rate of total calls carried in cell j
 R_{sh} = the rate of successful handoffs

The pdf of channel holding time T in a cell is derived as

$$f_T(t) = \left(\frac{\mu_C + \mu_{ni}}{1 + G} \right) e^{-(\mu_C + \mu_{ni})t} + \left(\frac{\mu_C + \mu_h}{1 + G} \right) e^{-(\mu_C + \mu_h)t} \quad (1.4)$$

where

$1/\mu_{ni}$ = the average channel holding time in cell i for a originating call
 $1/\mu_h$ = the average channel holding time for a handoff call
 G = the ratio of the offered rate of handoff requests to that of originating calls

1.5.1.3 Steele and Nofal's Traffic Model (Two-Dimensional)

Steele and Nofal [7] studied a traffic model based on city street microcells, catering to pedestrians making calls while walking along a street. From their assumptions, they showed that the arrival rate of handoff calls is

$$\lambda_H = \sum_{m=1}^6 [\lambda_O(1 - B_O) P_h \beta + \lambda_h(1 - P_f) P_{hh} \beta] \quad (1.5)$$

where

β = the fraction of handoff calls to the current cell from the adjacent cells
 $\lambda_h = 3\lambda_O(1 - B_O) P_I \beta$
 P_I = the probability that a new call that is not blocked will require at least one handoff

The average channel holding time T in a cell is

$$\bar{T} = \frac{(1 + \alpha_1)(1 - \gamma)}{\mu_w + \mu_C} + \frac{\gamma(1 + \alpha_2)}{\mu_o + \mu_C} + \frac{\alpha_1(1 - \gamma) + \gamma\alpha_2}{\mu_d + \mu_C} \quad (1.6)$$

where

$1/\mu_w$ = the average walking time of a pedestrian from the onset of the call until he reaches the boundary of the cell

$1/\mu_d$ = the average delay time a pedestrian spends waiting at the intersection to cross the road

$1/\mu_o$ = the average walking time of a pedestrian in the new cell

$$\alpha_1 = \mu_w P_{delay} / (\mu_d - \mu_w)$$

$$\alpha_2 = \mu_o P_{delay} / (\mu_d - \mu_o)$$

$P_{delay} = P_{cross} P_d$, the proportion of pedestrians leaving the cell by crossing the road

P_d = the probability that a pedestrian would be delayed when he crosses the road

$$\gamma = \lambda_H (1 - P_f') / [\lambda_H (1 - P_f') + \lambda_O (1 - B_O)]$$

1.5.1.4 Xie and Kuek's Traffic Model (One- and Two-Dimensional)

This model assumes a uniform density of mobile users throughout an area and that a user is equally likely to move in any direction with respect to the cell border. From this assumption, Xie and Kuek [8] showed that the arrival rate of handoff calls is

$$\lambda_H = E[C] \mu_{c-dwell} \quad (1.7)$$

where

$E[C]$ = the average number of calls in a cell

$\mu_{c-dwell}$ = the outgoing rate of mobile users.

The average channel holding time T in a cell is

$$\bar{T} = \frac{1}{\mu_C + \mu_{c-dwell}} \quad (1.8)$$

1.5.1.5 Zeng et al.'s Approximated Traffic Model (Any Dimensional)

Zeng et al.'s model is based on Xie and Kuek's traffic model [9]. Using Little's formula, when the blocking probability of originating calls and the forced termination probability of handoff calls are small, the average numbers of occupied channels $E[C]$ is approximated by

$$E[C] \approx \frac{\lambda_O + \lambda_H}{\mu} \quad (1.9)$$

where $1/\mu$ is the average channel holding time in a cell.

Therefore, the arrival rate of handoff calls is

$$\lambda_H \approx \frac{\mu_{c-dwell}}{\mu_C} \lambda_O \quad (1.10)$$

Xie and Kuek focused on the pdf of the speed of cell-crossing mobiles and refined previous results by making use of biased sampling. The distribution of mobile speeds of handoff calls used in Hong and Rappaport's traffic model has been adjusted by using

$$f^*(v) = \frac{vf(v)}{E[V]} \tag{1.11}$$

where $f(v)$ is the pdf of the random variable V (speed of mobile users), and $E[V]$ is the average of the random variable V .

$f^*(v)$ leads to the conclusion that the probability of handoff in Hong and Rappaport's traffic model is a pessimistic one, because the speed distribution of handoff calls are not the same as the overall speed distribution of all mobile users.

Steele's traffic model is not adaptive for an irregular cell and vehicular users. In Zeng et al.'s approximated traffic model, actual deviation from Xie and Kuek's traffic model is relatively small when the blocking probability of originating calls and the forced termination probability of handoff calls are small.

1.5.2 Handoff Schemes in Single Traffic Systems

In this section, we introduce nonpriority, priority, and queuing handoff schemes for a single traffic system such as either a voice or a data system [6–14]. Before introducing these schemes, we assume that a system has many cells, with each having S channels. The channel holding time has an exponential distribution with mean rate μ . Both originating and handoff calls are generated in a cell according to Poisson processes, with mean rates λ_O and λ_H , respectively. We assume the system with a homogeneous cell. We focus our attention on a single cell (called the marked cell). Newly generated calls in the marked cell are labeled originating calls (or new calls). A handoff request is generated in the marked cell when a channel holding MS approaches the marked cell from a neighboring cell with a signal strength below the handoff threshold.

1.5.2.1 Nonpriority Scheme

In this scheme, all S channels are shared by both originating and handoff request calls. The BS handles a handoff request exactly in the same way as an originating call. Both kinds of requests are blocked if no free channel is available. The system model is shown in Figure 1.3.

With the blocking call cleared (BCC) policy, we can describe the behavior of a cell as a $(S + 1)$ states Markov process. Each state is labeled by an integer i ($i = 0, 1, \dots, S$), repre-

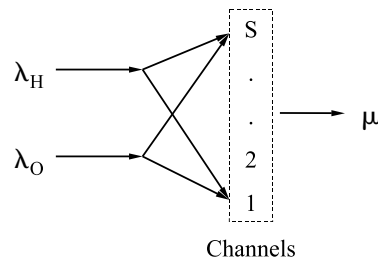


Figure 1.3 A generic system model for handoff.

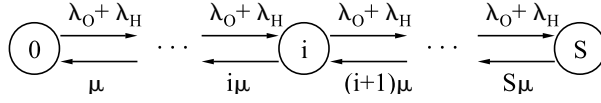


Figure 1.4 State transition diagram for Figure 1.3.

senting the number of channels in use. The state transition diagram is shown in Figure 1.4. The system model is modeled by a typical $M/M/S/S$ queueing model.

Let $P(i)$ be the probability that the system is in state i . The probabilities $P(i)$ can be determined in the usual way for birth–death processes. From Figure 1.4, the state equilibrium equation is

$$P(i) = \frac{\lambda_O + \lambda_H}{i\mu} P(i - 1), \quad 0 \leq i \leq S \tag{1.12}$$

Using the above equation recursively, along with the normalization condition

$$\sum_{i=0}^S P(i) = 1 \tag{1.13}$$

the steady-state probability $P(i)$ is easily found as follows:

$$P(i) = \frac{(\lambda_O + \lambda_H)^i}{i! \mu^i} P(0), \quad 0 \leq i \leq S \tag{1.14}$$

where

$$P(0) = \frac{1}{\sum_{i=0}^S \frac{(\lambda_O + \lambda_H)^i}{i! \mu^i}} \tag{1.15}$$

The blocking probability B_O for an originating call is

$$B_O = P(S) = \frac{\frac{(\lambda_O + \lambda_H)^S}{S! \mu^S}}{\sum_{i=0}^S \frac{(\lambda_O + \lambda_H)^i}{i! \mu^i}} \tag{1.16}$$

The blocking probability B_H of a handoff request is

$$B_H = B_O \tag{1.17}$$

Equation (1.16) is known as the Erlang-B formula.

A blocked handoff request call can still maintain the communication via current BS until the received signal strength goes below the receiver threshold or until the conversation is completed before the received signal strength goes below the receiver threshold.

1.5.2.2 Priority Scheme

In this scheme, priority is given to handoff requests by assigning S_R channels exclusively for handoff calls among the S channels in a cell. The remaining $S_C (= S - S_R)$ channels are shared by both originating calls and handoff requests. An originating call is blocked if the number of available channels in the cell is less than or equal to $S_R (= S - S_C)$. A handoff request is blocked if no channel is available in the target cell. The system model is shown in Figure 1.5.

We define the state i ($i = 0, 1, \dots, S$) of a cell as the number of calls in progress for the BS of that cell. Let $P(i)$ represent the steady-state probability that the BS is in state i . The probabilities $P(i)$ can be determined in the usual way for birth–death processes. The pertinent state transition diagram is shown in Figure 1.6. From the figure, the state balance equations are

$$\begin{cases} i\mu P(i) = (\lambda_O + \lambda_H) P(i-1) & 0 \leq i \leq S_C \\ i\mu P(i) = \lambda_H P(i-1) & S_C < i \leq S \end{cases} \quad (1.18)$$

Using this equation recursively, along with the normalization condition

$$\sum_{i=0}^S P(i) = 1 \quad (1.19)$$

the steady-state probability $P(i)$ is easily found as follows:

$$P(i) = \begin{cases} \frac{(\lambda_O + \lambda_H)^i}{i! \mu^i} P(0) & 0 \leq i \leq S_C \\ \frac{(\lambda_O + \lambda_H)^{S_C} \lambda_H^{i-S_C}}{i! \mu^i} P(0) & S_C \leq i \leq S \end{cases} \quad (1.20)$$

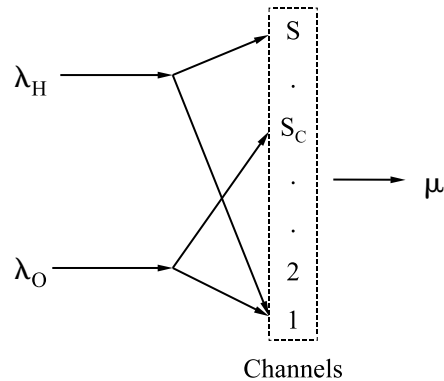


Figure 1.5 System model with priority for handoff call.

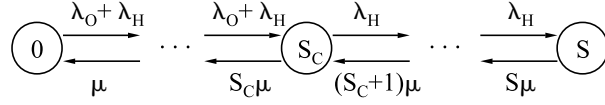


Figure 1.6 State transition diagram for Figure 1.5.

where

$$P(0) = \left[\sum_{i=0}^{S_C} \frac{(\lambda_O + \lambda_H)^i}{i! \mu^i} + \sum_{i=S_C+1}^S \frac{(\lambda_O + \lambda_H)^{S_C} \lambda_H^{i-S_C}}{i! \mu^i} \right]^{-1} \tag{1.21}$$

The blocking probability B_O for an originating call is given by

$$B_O = \sum_{i=S_C}^S P(i) \tag{1.22}$$

The blocking probability B_H of a handoff request is

$$B_H = P(S) = \frac{(\lambda_O + \lambda_H)^{S_C} \lambda_H^{S-S_C}}{S! \mu^S} P(0) \tag{1.23}$$

Here again, a blocked handoff request call can still maintain the communication via current BS until the received signal strength goes below the receiver threshold or the conversation is completed before the received signal strength goes below the receiver threshold.

1.5.2.3 Handoff Call Queuing Scheme

This scheme is based on the fact that adjacent cells in a mobile cellular radio system are overlaid. Thus, there is a considerable area (i.e., handoff area) where a call can be handled by BSs in adjacent cells. The time a mobile user spent moving across the handoff area is referred as the degradation interval. In this scheme, we assume that the same channel sharing scheme is used as that of a priority scheme, except that queuing of handoff requests is allowed. The system model is shown in Figure 1.7.

To analyze this scheme, it is necessary to consider the handoff procedure in more detail. When a MS moves away from the BS, the received signal strength decreases, and when it gets lower than a threshold level, the handoff procedure is initiated. The handoff area is defined as the area in which the average received signal strength of a MS receiver from the BS is between the handoff threshold level and the receiver threshold level.

If the BS finds all channels in the target cell occupied, a handoff request is put in the queue. If a channel is released when the queue for handoff requests is not empty, the channel is assigned to request on the top of the queue. If the received signal strength from the current BS falls below the receiver threshold level prior to the mobile being assigned a channel in the target cell, the call is forced to termination. The first-in-first-out

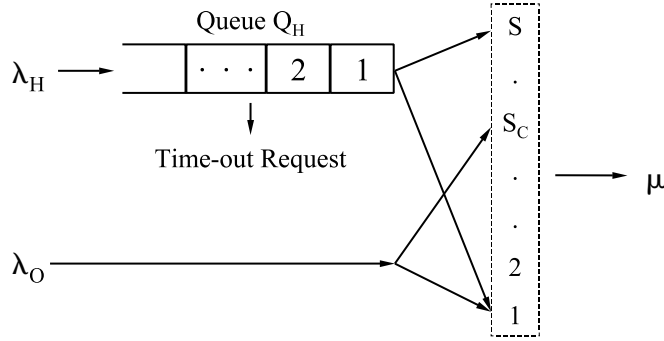


Figure 1.7 System model with priority and queue for handoff call.

(FIFO) queueing strategy is used and infinite queue size at the BS is assumed. For a finite queue size, see the discussion in the next section. The duration of a MS in the handoff area depends on system parameters such as the moving speed, the direction of the MS, and the cell size. We define this as the dwell time of a mobile in the handoff area and denote it by random variable $T_{h-dwell}$. For simplicity of analysis, we assume that this dwell time is exponentially distributed with mean $E[T_{h-dwell}] (= 1/\mu_{h-dwell})$.

Let us define the state i ($i = 0, 1, \dots, \infty$) of a cell as the sum of channels being used in the cell and the number of handoff requests in the queue. It is apparent from the above assumptions that i is a one-dimensional Markov chain. The state transition diagram of the cell is given in Figure 1.8. The equilibrium probabilities $P(i)$ are related to each other through the following state balance equations:

$$\begin{cases} i\mu P(i) = (\lambda_O + \lambda_H)P(i-1) & 0 \leq i \leq S_C \\ i\mu P(i) = \lambda_H P(i-1) & S_C < i \leq S \\ [S\mu + (i-S)(\mu_C + \mu_{h-dwell})] P(i) = \lambda_H P(i-1) & S < i \leq \infty \end{cases} \quad (1.24)$$

Using the above equation recursively, along with the normalization condition of equation (1.13), the steady-state probability $P(i)$ is easily found as follows:

$$P(i) = \begin{cases} \frac{(\lambda_O + \lambda_H)^i}{i! \mu^i} P(0) & 0 \leq i \leq S_C \\ \frac{(\lambda_O + \lambda_H)^{S_C} \lambda_H^{i-S_C}}{i! \mu^i} P(0) & S_C < i \leq S \\ \frac{(\lambda_O + \lambda_H)^{S_C}}{S! \mu^S} \frac{\lambda_H^{i-S_C}}{\prod_{j=1}^{i-S} [S\mu + j(\mu_C + \mu_{h-dwell})]} P(0) & S < i \leq \infty \end{cases} \quad (1.25)$$

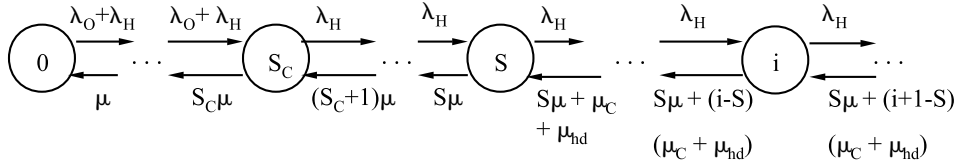


Figure 1.8 State transition diagram for Figure 1.7.

where

$$P(0) = \left\{ \sum_{i=0}^{S_C} \frac{(\lambda_O + \lambda_H)^i}{i! \mu^i} + \sum_{i=S_C+1}^S \frac{(\lambda_O + \lambda_H)^{S_C} \lambda_H^{i-S_C}}{i! \mu^i} + \sum_{i=S+1}^{\infty} \frac{(\lambda_O + \lambda_H)^{S_C}}{S! \mu^S} \frac{\lambda_H^{i-S_C}}{\prod_{j=1}^{i-S} [S\mu + j(\mu_C + \mu_{h-dwell})]} \right\}^{-1} \quad (1.26)$$

The blocking probability B_O for an originating call is

$$B_O = \sum_{i=S_C}^S P(i) \quad (1.27)$$

The forced termination probability B_H of handoff requests is

$$P_f' = \sum_{k=0}^{\infty} P(S+k) P_{f|k} \quad (1.28)$$

where $P_{f|k}$ is a probability that a handoff request fails after joining the queue in position $k + 1$ and it is given by

$$P_{f|k} = 1 - \left(\frac{\mu_C + \mu_{h-dwell}}{\mu S + \mu_C + \mu_{h-dwell}} \right) \prod_{i=1}^k \left\{ 1 - \left(\frac{\mu_C + \mu_{h-dwell}}{\mu S + \mu_C + \mu_{h-dwell}} \right) \left(\frac{1}{2} \right)^i \right\} \quad (1.29)$$

This scheme can be said to be equivalent to the $S_C = S$.

1.5.2.4 Originating and Handoff Calls Queuing Scheme

We consider a system with many cells, each having S channels. In the BS, there are two queues Q_H and Q_O for handoff requests and originating calls, respectively (Figure 1.9). The capacities of Q_H and Q_O are M_H and M_O , respectively. A handoff request is queued in Q_H if it finds no free channels on arrival. On the other hand, an originating call is queued in Q_O when on arrival it finds available channels less than or equal to $(S - S_C)$. A request call is blocked if on arrival its own queue is full.

An originating call in the queue is deleted from the queue when it moves out of the cell before getting a channel. Also, a handoff request is deleted from the queue when it passes

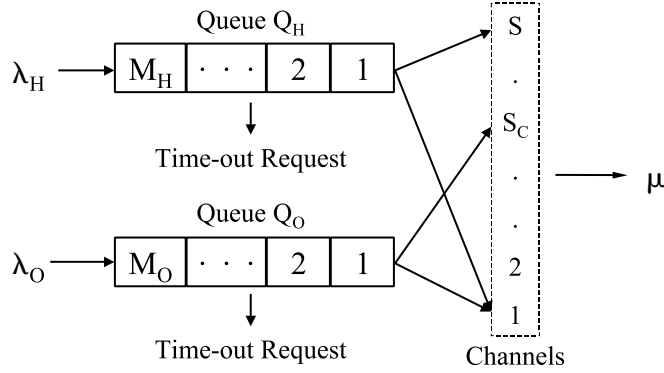


Figure 1.9 System model.

through the handoff area before getting a new channel (i.e., forced termination) or the conversation is completed before passing through the handoff area. A blocked handoff request call can still maintain the communication via the current BS until the received signal strength goes below the receiver threshold or the conversation is completed before passing through the handoff area. A blocked handoff call can repeat trial handoffs until the received signal strength goes below the receiver threshold. However, the capacity of M_H of queue Q_H is usually large enough so that the blocking probability of handoff request calls can be neglected. Thus, repeated handoff requests are excluded from any discussion.

In this method, the state of the marked cell is defined by a two-tuple of nonnegative integers (i, j) , where i is the sum of s_b busy channels and j is the number of originating calls in Q_O . Note that if $s_b < S$, then $i = s_b$, and $i = q_h + S$ when $s_b = S$, where q_h is the number of handoff requests in Q_H . It is apparent from the above assumptions that (i, j) is a two-dimensional Markov chain. The state transition diagram of the cell is given in Figure 1.10.

Since the sum of all state probabilities $P(i, j)$ is equal to 1, we have

$$\sum_{i=0}^{S_C-1} P(i, 0) + \sum_{i=S_C}^{S+M_H} \sum_{j=0}^{M_O} P(i, j) = 1 \quad (1.30)$$

In the state transition diagram there are $N_T = (S + M_H + 1)(M_O + 1) - S_C M_O$ states. Therefore, there are N_T balance equations. However, note that any one of these balance equations can be obtained from other $N_T - 1$ equations. Adding the normalizing equation (1.30), we can obtain N_T independent equations. Though N_T is usually rather large, we can obtain all the probabilities $P(i, j)$ (for $i = 0, 1, 2, \dots, S + M_H$ and $j = 0, 1, 2, \dots, M_O$) using the following iterative method.

Step 1: Select an arbitrary initial (positive) value for λ_H . [If we use λ_H given by (1.10), we can improve the speed of the convergence.]

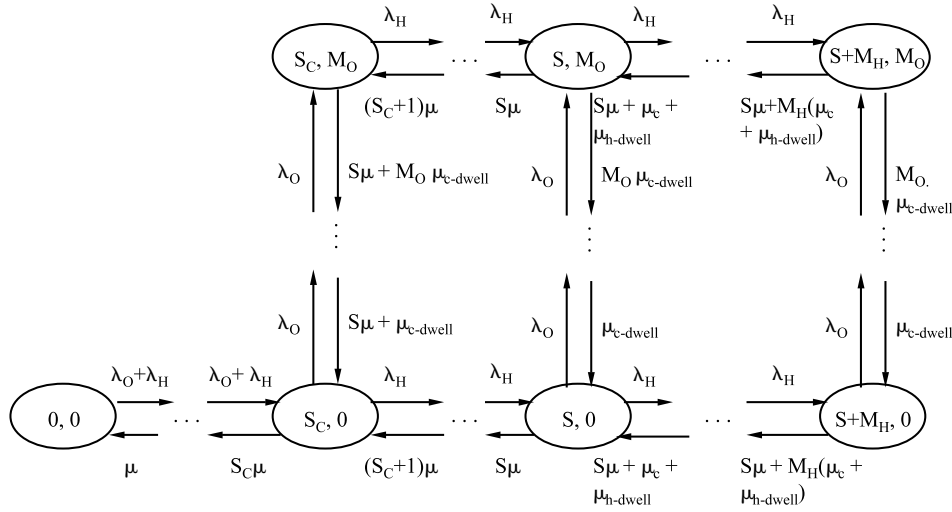


Figure 1.10 State transition diagram.

Step 2: Compute all the probabilities $P(i, j)$ (for $i = 0, 1, 2, \dots, S + M_H$ and $j = 0, 1, 2, \dots, M_O$) using SOR (successive over-relation) method.

Step 3: Compute the average number of calls holding channels using the following formula:

$$E[C] = \sum_{i=1}^{S_C-1} iP(i, 0) + \sum_{i=S_C}^S i \sum_{j=0}^{M_O} P(i, j) \quad (1.31)$$

Step 4: Compute new λ_H substituting (1.31) into (1.7). If $|new \lambda_H - old \lambda_H| \leq \epsilon$, stop execution. Otherwise, go to Step 2. Here ϵ is a small positive number to check the convergence.

Based on the above $P(i, j)$ s, we can obtain the following performance measures of the system.

The blocking probability B_O of an originating call is

$$B_O = \sum_{i=S_C}^{S+M_H} P(i, M_O) \quad (1.32)$$

The blocking probability B_H of a handoff request is equal to the probability of its own queue being filled up. Thus,

$$B_H = \sum_{j=0}^{M_O} P(S + M_H, j) \quad (1.33)$$

16 HANDOFF IN WIRELESS MOBILE NETWORKS

The average L_O length of queue Q_O is:

$$L_O = \sum_{j=1}^{M_O} j \sum_{i=S_C}^{S+M_H} P(i, j) \quad (1.34)$$

and the average length L_N of queue Q_H is:

$$L_H = \sum_{i=S+1}^{S+M_H} (i - S) \sum_{j=0}^{M_O} P(i, j) \quad (1.35)$$

Since the average number of originating calls arrived and deleted from the queue in unit time are $(1 - B_O)\lambda_O$ and $\mu_{c-dwell}L_O$, respectively, the time-out probability of originating calls is given by

$$P_{O-out} = \frac{\mu_{c-dwell}L_O}{(1 - B_O)\lambda_O} \quad (1.36)$$

Similarly, the time-out probability of handoff request calls in the queue Q_H is given by

$$P_{H-out} = \frac{\mu_{h-dwell}L_H}{(1 - B_H)\lambda_H} \quad (1.37)$$

Therefore, the probability of an originating call not being assigned a channel and the forced termination probability of a handoff request are given by

$$P_O = B_O + (1 - B_O)P_{O-out} \quad (1.38)$$

and

$$P_f' = B_H + (1 - B_H)P_{H-out} \quad (1.39)$$

Once a MS is assigned to a channel and if a call is in progress, any subsequent cell boundary crossings necessitates further handoffs. The handoff probability P_h of a call is the probability that the call holding time T_C (random variable) exceeds the dwell time $T_{c-dwell}$ (random variable) of the user in a cell, i.e.,

$$P_h = \Pr\{T_C > T_{c-dwell}\} \quad (1.40)$$

Assuming that T_C and $T_{c-dwell}$ are independent, we can easily get

$$P_h = \frac{\mu_{c-dwell}}{\mu_C + \mu_{c-dwell}} \quad (1.41)$$

The forced termination probability P_f that a call accepted by the system is forced to terminate during its lifetime is a true measure of the system performance. It is important

to distinguish between this probability and the failure probability P'_f of a single handoff attempt. The forced termination probability P_f of handoff calls can be expressed as

$$P_f = \sum_{f=1}^{\infty} P_h P'_f [(1 - P'_f) P_h]^{f-1} = \frac{P_h P'_f}{1 - P_h(1 - P'_f)} \quad (1.42)$$

For special situations, solutions are already known for the case of $M_H = \infty$ and $M_H = 0$ when $M_O = 0$. In the system with $M_H = \text{finite}$, an originating call is blocked if the number of available channels in the cell is less than or equal to $S - S_C$. A handoff request is blocked if on arrival it finds that Q_H is filled.

In this case, we consider the case for $M_O = 0$. The two-dimensional state-transition diagram becomes one-dimensional ($j = 0$). Therefore, the state probabilities can easily be obtained as follows:

$$P(i,0) = \begin{cases} \frac{a^i}{i!} P(0,0) & 0 \leq i \leq S_C \\ \left(\frac{a}{b}\right)^{S_C} \frac{b^i}{i!} P(0,0) & S_C \leq i \leq S \\ \frac{\left(\frac{a}{b}\right)^{S_C}}{S!} \frac{b^i P(0,0)}{\prod_{j=1}^{i-S} (S+jh)} & S+1 \leq i \leq S+M_H \end{cases} \quad (1.43)$$

where $a = \frac{\lambda_O + \lambda_H}{\mu}$, $b = \frac{\lambda_H}{\mu}$, and $h = \frac{\mu_C + \mu_{h\text{-dwell}}}{\mu}$,

$$P(0,0) = \left\{ \sum_{i=0}^{S_C} \frac{a^i}{i!} + \left(\frac{a}{b}\right)^{S_C} \sum_{i=S_C+1}^S \frac{b^i}{i!} + \frac{\left(\frac{a}{b}\right)^{S_C}}{S!} \sum_{i=S+1}^{S+M_H} \frac{b^i}{\prod_{j=1}^{i-S} (S+jh)} \right\}^{-1} \quad (1.44)$$

and

$$\lambda_H = E[C] \mu_{c\text{-dwell}} = \mu_{c\text{-dwell}} \sum_{i=1}^S iP(i,0) \quad (1.45)$$

Therefore, the blocking probability B_O of an originating call is

$$B_O = \sum_{i=S_C}^{S+M_H} P(i,0) = \left[\frac{a^{S_C}}{S_C!} + \left(\frac{a}{b}\right)^{S_C} \sum_{i=S_C+1}^S \frac{b^i}{i!} + \frac{\left(\frac{a}{b}\right)^{S_C}}{S!} \sum_{i=S+1}^{S+M_H} \frac{b^i}{\prod_{j=1}^{i-S} (S+jh)} \right] P(0,0). \quad (1.46)$$

The average length L_H of queue Q_H is

$$L_H = \sum_{i=S+1}^{S+M_H} (i-S)P(i, 0). \quad (1.47)$$

1.5.3 Handoff Schemes in Multiple Traffic Systems

In this section, we discuss nonpreemptive and preemptive priority handoff schemes for a multiple traffic system, such as an integrated voice and data system or integrated real-time and nonreal-time system. Although we focus our attention just on integrated voice and data systems, the results can be extended to other similar systems. Before introducing these schemes, we make the following assumptions for our discussion. The call holding time T_{CV} of voice calls is assumed to have an exponential distribution with mean $E[T_{CV}] (= 1/\mu_{CV})$. The data length T_{CD} is also assumed to have an exponential distribution with mean $E[T_{CD}] (= 1/\mu_{CD})$. The dwell time $T_{c\text{-dwell}}$ (the random variable) of mobile users in a cell is assumed to have an exponential distribution with mean $E[T_{c\text{-dwell}}] (= 1/\mu_{c\text{-dwell}})$. The random variable $T_{h\text{-dwell}}$ is defined as the time spent in the handoff area by voice handoff request calls and is assumed to have an exponential distribution with mean $E[T_{h\text{-dwell}}] (= 1/\mu_{c\text{-dwell}})$. The channel holding time of a voice (or data) call is equal to the smaller one between $T_{c\text{-dwell}}$ and T_{CV} (or T_{CD}). Using the memoryless property of the exponential pdf, we see that the random variables T_V and T_D (the channel holding time of voice and data calls) are both exponentially distributed, with means $E[T_V] [= 1/\mu_V = 1/(\mu_{CV} + \mu_{c\text{-dwell}})]$ and $E[T_D] [= 1/\mu_D = 1/(\mu_{CD} + \mu_{c\text{-dwell}})]$, respectively. We assume that the arrival processes of originating voice and data calls and voice and data handoff calls in a cell are Poisson. The arrival rates of originating voice and data calls are designated as λ_{OV} and λ_{OD} , respectively. We denote the arrival rates of voice and data handoff requests by λ_{HV} and λ_{HD} , respectively. A data handoff request in the queue of the current cell is transferred to the queue of target cell when it moves out of the cell before getting a channel. The transfer rate is given by

$$\lambda_{\text{time-out}} = L_{qd}\mu_{c\text{-dwell}} \quad (1.48)$$

where L_{qd} is the average length of data queue.

We define a new variable λ_{HT} by

$$\lambda_{HT} = \lambda_{HD} + \lambda_{\text{time-out}} = N_D\mu_{c\text{-dwell}} \quad (1.49)$$

where N_D is the average number of data handoff requests in a cell.

1.5.3.1 Nonpreemptive Priority Handoff Scheme

We consider a system with many cells each having S channels. As the system is assumed to have homogeneous cells, we focus our attention on a single cell called the marked cell. A system model is shown in Figure 1.11. In each BS, there are two queues, Q_V and Q_D , with capacities M_V and M_D for voice and data handoff requests, respectively.

Newly generated calls in the marked cell are called originating calls. For voice users, there is a handoff area. For data users, the boundary is defined as the locus of points where

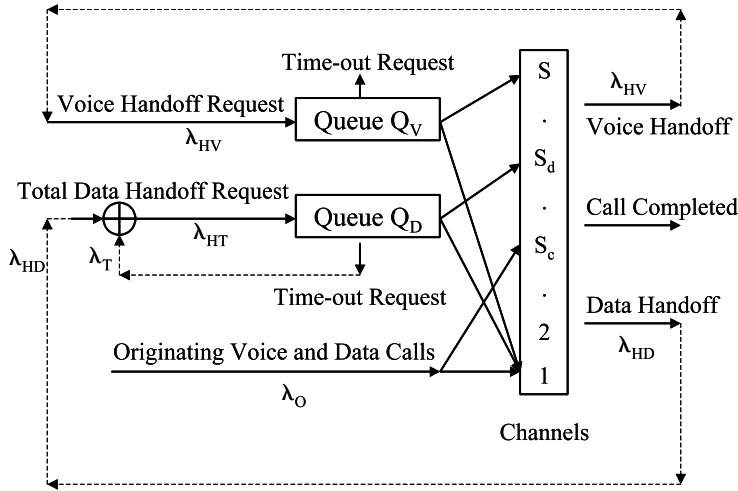


Figure 1.11 System model with two queues for handoffs.

the average received signal strength of the two neighboring cells are equal. The process of generation for handoff request is same as in previous schemes.

A voice handoff request is queued in Q_V on arrival if it finds no idle channels. On the other hand, a data handoff request is queued in Q_D on arrival when it finds $(S - S_d)$ or fewer available channels, where S_d is the number of usable channels for data handoff users. An originating voice or an originating data call is blocked on arrival if it finds $(S - S_c)$ or fewer available channels, where S_c is the number of channels for both originating calls. No queue is assumed here for originating calls. A handoff request is blocked if its own queue is full on its arrival.

If there are channels available, the voice handoff request calls in Q_V are served based on the FIFO rule. If more than $(S - S_d)$ channels are free, the data handoff request calls in Q_D are served by the FIFO rule. A voice handoff request in the queue is deleted from the queue when it passes through the handoff area before getting a new channel (i.e., forced termination) or its communication is completed before passing through the handoff area. A data handoff request can be transferred from the queue of the current cell to the one of the target cells when it moves out of the current cell before getting a channel.

A blocked voice handoff request maintains communication via the current BS until the received signal strength falls below the receiver threshold or the conversation is completed before passing through the handoff area. However, the probability of a blocked voice handoff request call completing the communication before passing through the handoff area is neglected. A blocked voice handoff request can repeat trial handoffs until the received signal strength goes below the receiver threshold. However, the capacities of M_V and M_D of queues are usually large enough so that the blocking probability of handoff request calls can be neglected. Thus, repeated trials of blocked handoff requests are excluded from any discussion.

The state of the marked cell is defined by a three-tuple of nonnegative integers (i, j, k) ,

20 HANDOFF IN WIRELESS MOBILE NETWORKS

where i is the sum of the number of channels used by voice calls (including originating calls and handoff requests) and the number of voice handoff requests in the queue Q_V , j is the number of channels used by data handoff requests, and k is the number of data handoff requests in the queue Q_D . It is apparent from the above assumptions that (i, j, k) is a three-dimensional Markov chain.

In the state transition diagram, there are $N_T = (S + M_V - S_d + 1)(S_d + 1)(M_D + 1) + S_d(S_d + 1)/2$ states. Therefore, the state transition diagram leads to N_T balance equations. Equilibrium probabilities $P(i, j, k)$ are related to each other through the state balance equations. However, note that any one of these balance equations can be obtained from other $N_T - 1$ equations. Since the sum of all state probabilities is equal to 1, we have

$$\sum_{j=0}^{S_d} \sum_{i=0}^{S+M_V-j} P(i, j, 0) + \sum_{j=0}^{S_d} \sum_{i=S_d-j}^{S+M_V-j} \sum_{k=1}^{M_D} P(i, j, k) = 1 \quad (1.50)$$

Adding the normalizing equation (1.50), we can obtain N_T independent equations in which λ_{HV} and λ_{HT} are two unknown variables. Using equation (1.7), we can get

$$\lambda_{HV} = E[C_V] \mu_{c-dwell} \quad (1.51)$$

Adding equations (1.49) and (1.51) leads to $N_T + 2$ nonlinear independent simultaneous equations. Though N_T is usually rather large, all the probabilities $P(i, j, k)$ (for $i = 0, 1, 2, \dots, S + M_V$; $j = 0, 1, 2, \dots, S_d$, and $k = 0, 1, 2, \dots, M_D$) can be obtained by solving $N_T + 2$ nonlinear independent simultaneous equations, as illustrated in the next section.

Step 1: Select arbitrary initial (positive) values for λ_{HV} and λ_{HT} .

Step 2: Compute all the probabilities $P(i, j, k)$ (for $i = 0, 1, 2, \dots, S + M_V$, $j = 0, 1, 2, \dots, S_d$, and $k = 0, 1, 2, \dots, M_D$) using the SOR method.

Step 3: Compute the average numbers $E[C_V]$ of voice calls holding channels, and compute the average numbers N_D of data channel requests in a cell using the following relations:

$$\begin{aligned} E[C_V] &= \sum_{i=0}^{S-S_d} i \sum_{j=0}^{S_d} \sum_{k=0}^{M_D} P(i, j, k) + \sum_{i=S-S_d+1}^S i \sum_{j=0}^{S-i} \sum_{k=0}^{M_D} P(i, j, k) \\ &+ \sum_{j=0}^{S_d} (S-j) \sum_{i=S_d-j+1}^{S+M_V-j} \sum_{k=0}^{M_D} P(i, j, k) \end{aligned} \quad (1.52)$$

and

$$N_D = \sum_{j=1}^{S_d} j \sum_{i=0}^{S+M_V-j} \sum_{k=0}^{M_D} P(i, j, k) + \sum_{k=1}^{M_D} k \sum_{j=0}^{S_d} \sum_{i=S_d-j}^{S+M_V-j} P(i, j, k) \quad (1.53)$$

Step 4: Compute new λ_{HV} by substituting (1.52) into (1.51). Compute new λ_{HT} by substituting (1.53) into (1.49). If $|new \lambda_{HV} - old \lambda_{HV}| \leq \varepsilon$ and $|new \lambda_{HT} - old \lambda_{HT}| \leq \varepsilon$,

stop execution. Otherwise, go to Step 2. Here ε is a small positive number to check the convergence.

Based on the above $P(i, j, k)$ s, we can obtain the following performance measures of the system.

The blocking probability of an originating voice call or originating data call is

$$B_O = B_{OV} = B_{OD} = \sum_{j=0}^{S_d} \sum_{i=S_d-j}^{S+M_V-j} \sum_{k=0}^{M_D} P(i, j, k) \quad (1.54)$$

The blocking probability B_{HV} of a voice handoff request is

$$B_{HV} = \sum_{j=0}^{S_d} \sum_{k=0}^{M_D} P(S + M_V - j, j, k) \quad (1.55)$$

The blocking probability B_{HD} of a data handoff request is

$$B_{HD} = \sum_{j=0}^{S_d} \sum_{i=S_d-j}^{S+M_V-j} P(i, j, M_D) \quad (1.56)$$

The average length L_{qv} of queue Q_V is

$$L_{qv} = \sum_{i=S+1}^{S+M_V} (i - S) \sum_{j=0}^{S_d} \sum_{k=0}^{M_D} P(i - j, j, k) \quad (1.57)$$

The average length L_{qd} of queue Q_D is

$$L_{qd} = \sum_{k=1}^{M_D} k \sum_{j=0}^{S_d} \sum_{i=S_d-j}^{S+M_V-j} P(i, j, k) \quad (1.58)$$

Since the average number of voice handoff requests arrived and deleted in unit time are $(1 - B_{HV})\lambda_{HV}$ and $\mu_{h-dwell}L_{qv}$, respectively, the time-out probability of a voice handoff requests in the queue Q_V is given by

$$P_{V-out} = \frac{\mu_{h-dwell}L_{qv}}{(1 - B_{HV})\lambda_{HV}} \quad (1.59)$$

Therefore, failure probability of a voice handoff request for single handoff attempt is

$$P'_{jV} = B_{HV} + (1 - B_{HV}) P_{V-out} \quad (1.60)$$

The voice user in a cell is given by

$$P_h = \Pr\{T_{CV} > T_{c-dwell}\} \quad (1.61)$$

22 HANDOFF IN WIRELESS MOBILE NETWORKS

Assuming that T_{CV} and $T_{c-dwell}$ are independent, we can easily get

$$P_h = \frac{\mu_{c-dwell}}{\mu_{CV} + \mu_{c-dwell}} \quad (1.62)$$

The forced termination probability P_{fV} of voice calls can be expressed as

$$P_{fV} = \sum_{l=1}^{\infty} P_h P'_{fV} [(1 - P'_{fV}) P_h]^{l-1} = \frac{P_h P'_{fV}}{1 - P_h (1 - P'_{fV})} \quad (1.63)$$

Using Little's formula, the average value of waiting time T_W of data handoff requests in the queue is given by

$$T_W = \frac{L_{qd}}{(1 - B_{OD}) \lambda_{OD} + (1 - B_{HD}) \lambda_{HT}} \quad (1.64)$$

Average value of time T_S (random variable) of a call in a cell is

$$T_S = \frac{N_D}{(1 - B_{OD}) \lambda_{OD} + (1 - B_{HD}) \lambda_{HT}} \quad (1.65)$$

Let us define N_h as the average number of handoffs per data handoff request during its lifetime. Thus, we have

$$N_h = \frac{N_h T_W + E[T_{CD}]}{T_S} \quad (1.66)$$

Then,

$$N_h = \frac{(1 - B_{OD}) \lambda_{OD} + (1 - B_{HD}) \lambda_{HT}}{E[C_D] \mu_{CD}} \quad (1.67)$$

where

$$E[C_D] = \sum_{j=1}^{S_d} j \sum_{i=0}^{S+M_V-j} \sum_{k=0}^{M_D} P(i, j, k) \quad (1.68)$$

Therefore, the average transmission delay (except average data length) T_{delay} of data is

$$T_{\text{delay}} = N_h T_W = \frac{L_{qd}}{E[C_D] \mu_{CD}} \quad (1.69)$$

1.5.3.2 Preemptive Priority Handoff Scheme

This scheme is a modification of a nonpreemptive priority handoff scheme, with higher priorities for voice handoff request calls. In this scheme, a handoff request call is served if there are channels available when such a voice handoff request call arrives. Otherwise, the voice handoff request can preempt the data call, when we assume there is an ongoing data call, if on arrival it finds no idle channel. The interrupted data call is returned to the data queue Q_D and waits for a channel to be available based on the FIFO rule. A voice handoff request is queued in Q_V by the system if all the channels are occupied by prior calls and the data queue Q_D is full (i.e., data calls cannot be preempted by voice handoff calls when the data queue Q_D is full). It is possible to think of another scheme where data calls in service can be preempted by voice handoff calls irrespective of whether the queue Q_D is full or not. However, the same effect can be observed if the queue capacity is increased to a relatively large value.

The same state of the marked cell is assumed and represented by a three-tuple of non-negative integers (i, j, k) as defined in the nonpreemptive priority handoff scheme. In the state transition diagram for the three-dimensional Markov chain model there are $N_T = (S - S_d + 1)(S_d + 1)(M_D + 1) + (S_d + M_D + 1)M_V + S_d(S_d + 1)/2$ states. Therefore, as in the nonpreemptive priority handoff scheme, we can get N_T balance equations through the state transition diagram. Equilibrium probabilities $P(i, j, k)$ are related to each other through the state balance equations. However, note that any one of these balance equations can be obtained from other $N_T - 1$ equations. Since the sum of all state probabilities is equal to 1, we have

$$\begin{aligned} & \sum_{j=0}^{S_d} \sum_{i=0}^{S-j} P(i, j, 0) + \sum_{j=0}^{S_d} \sum_{i=S_d-j}^{S-j} \sum_{k=1}^{M_D} P(i, j, k) + \sum_{i=S+1}^{S+M_V} \sum_{k=0}^{M_D} P(i, 0, k) \\ & + \sum_{j=1}^{S_d} \sum_{i=S-j+1}^{S+M_V-j} P(i, j, M_D) = 1 \end{aligned} \quad (1.70)$$

The probabilities $P(j, j, k)$ (for $i = 0, 1, 2, \dots, S + M_V; j = 0, 1, 2, \dots, S_d$, and $k = 0, 1, 2, \dots, M_D$) can be obtained by using the same method of computation in the nonpreemptive priority handoff scheme. The differences are:

$$\begin{aligned} E[C_V] &= \sum_{i=0}^{S-S_d} i \sum_{j=0}^{S_d} \sum_{k=0}^{M_D} P(i, j, k) + \sum_{i=S-S_d+1}^S i \sum_{j=0}^{S-i} \sum_{k=0}^{M_D} P(i, j, k) \\ & + \sum_{j=0}^{S_d} (S-j) \sum_{i=S-j+1}^{S+M_V-j} P(i, j, M_D) \end{aligned} \quad (1.71)$$

$$\begin{aligned} N_D &= \sum_{j=1}^{S_d} j \sum_{i=0}^{S-j} \sum_{k=0}^{M_D} P(i, j, k) + \sum_{j=1}^{S_d} j \sum_{i=S-j+1}^{S+M_V-j} P(i, j, M_D) + \sum_{k=1}^{M_D} k \sum_{j=0}^{S_d} \sum_{i=S_d-j}^{S-j} P(i, j, k) \\ & + \sum_{k=1}^{M_D-1} k \sum_{i=S+1}^{S+M_V} P(i, 0, k) + M_D \sum_{j=0}^{S_d} \sum_{i=S-j+1}^{S+M_V-j} P(i, j, M_D) \end{aligned} \quad (1.72)$$

24 HANDOFF IN WIRELESS MOBILE NETWORKS

and

$$E[C_D] = \sum_{j=1}^{S_d} j \sum_{i=0}^{S-j} \sum_{k=0}^{M_D} P(i, j, k) + \sum_{j=1}^{S_d} j \sum_{i=S-j+1}^{S+M_V-j} P(i, j, M_D) \quad (1.73)$$

Therefore, the performance measurements can be obtained by equations (1.54)–(1.69).

1.6 SUMMARY

The basic concept of handoff in mobile cellular radio systems has been introduced. Several different traffic models have been described and briefly discussed. Four conventional handoff schemes in single traffic systems—i.e., nonpriority scheme, priority scheme, handoff call queueing scheme, and originating and handoff call queueing schemes—have been summarized in this chapter. The two handoff schemes with and without preemptive priority procedures for integrated voice and data wireless mobile networks have also been covered in detail.

REFERENCES

1. M. Gudmundson, Analysis of handover algorithms, *Proc. IEEE VTC '91*, pp. 537–542, May 1991.
2. V. Kapoor, G. Edwards, and R. Snkar, Handoff criteria for personal communication networks, *Proc. IEEE ICC '94*, pp. 1297–1301, May 1994.
3. G. P. Pollini, Trends in handover design, *IEEE Commun. Magazine*, pp. 82–90, March 1996.
4. N. D. Tripathi, J. H. Reed, and H. F. Vanlandingham, Handoff in Cellular Systems, *IEEE Personal Commun.*, December 1998.
5. D. Hong and S. S. Rappaport, Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures, *IEEE Trans. Veh. Technol.*, Vol. VT-35, No. 3, pp. 448–461, August 1986.
6. S. A. El-Dolil, W. C. Wong, and R. Steele, Teletraffic performance of highway microcells with overlay macrocell, *IEEE J. Select. Areas in Commun.*, Vol. 7, No. 1, pp. 71–78, January 1989.
7. R. Steele and M. Nofal, Teletraffic performance of microcellular personal communication networks, *IEEE PROCEEDINGS-I*, Vol. 139, No. 4, August 1992.
8. H. Xie and S. Kuek Priority handoff analysis, *Proc. IEEE VTC '93*, pp. 855–858, 1993.
9. Q-A. Zeng, K. Mukumoto, and A. Fukuda, Performance analysis of mobile cellular radio systems with two-level priority reservation handoff procedure, *IEICE Trans. Commun.*, Vol. E80-B, No. 4, pp. 598–604, April 1997.
10. S. Tekinay and B. Jabbari, A measurement-based prioritization scheme for handovers in mobile cellular networks, *IEEE J. Select. Areas in Commun.*, Vol. 10, No. 8, Oct. 1992.
11. Q-A. Zeng, K. Mukumoto, and A. Fukuda, Performance analysis of mobile cellular radio systems with priority reservation handoff procedures, *Proc. IEEE VTC '94*, Vol. 3, pp. 1829–1833, June 1994.

12. R. B. Cooper, *Introduction to Queueing Theory*, 2nd ed. New York: Elsevier North Holland, 1981.
13. J. D. Wells, Cellular system design using the expansion cell layout method, *IEEE Trans. Veh. Technol.*, Vol. VT-33, May 1984.
14. H. Akimaru and R. B. Cooper, *Teletraffic Engineering*. Ohm, 1985.
15. Q-A. Zeng and D. P. Agrawal, Performance analysis of a handoff scheme in integrated voice/data wireless networks, *Proc. IEEE VTC 2000 Fall*, Vol. 4, pp. 1986–1992, September 2000.
16. Q-A. Zeng and D. P. Agrawal, An analytical modeling of handoff for integrated voice/data wireless networks with priority reservation and preemptive priority procedures, *Proc. ICPP 2000 Workshop on Wireless Networks and Mobile Computing*, pp. 523–529, August 2000.

