CHAPTER 2

MOBILITY BASED, PARAMETERS MEASUREMENTS BASED AND FUZZY BASED CAC SCHEMES IN WCDMA SYSTEM

The wide-band CDMA (WCDMA) technology has emerged as the main air interface for 3G wireless systems, which promises to provide a transmission rate from 144kbps to 2Mbps, enabling multimedia services as those provided by broadband wired networks. The Radio Resource Management (RRM) module in the cellular network system is responsible for efficient utilization of air interface resources and guarantee a certain QoS level to different users according to their traffic profiles. The Call Admission Control (CAC) mechanism is one of the most important components of RRM. The radio resource-reservation estimation (RRE) mechanism helps CAC to decide the amount of resource (Signal to Interference Ratio Guard Margin) to be reserved in order to provide QoS guarantees to mobile users throughout the lifetime of the connection and also to resolve the inter-cell unbalanced traffic problem.

As it is impractical to completely eliminate handoff call dropping, the best one could do, is to keep the handoff dropping probability ($P_{hd}$) below a target level. Moreover, maximizing resource utilization while keeping $P_{nbr}$, the probability of new call blocking, below a target value, is another critical factor for evaluating CAC algorithms. Based on the above considerations, several schemes have recently been proposed for CAC in wireless cellular networks. The guard channel policy (Lu et al 1996) and fractional guard
channel policy (Naghshineh et al 1996) determine the number of guard channels for handoffs by considering just the status of the local cell. Users are assumed to be uniformly located in any cell of the mobile network under these policies. The distributed call admission control scheme (Papoulis 1991) considers not only the status of the local cell but also that of adjacent cells. The shadow cluster scheme (Lu et al 1996) estimates future resource requirements in a collection of cells in which a mobile is likely to visit in the future. Admission control is performed based on this estimate. However, this proposal lacks a mechanism to determine the shadow cluster in real networks, as it assumes either precise knowledge of user mobility or totally random user movements. There have also been some research efforts to predict user mobility. In the paper by Fang et al (1999), the next cell to which a mobile will move is predicted in an indoor environment. But this scheme does not estimate channel holding time and therefore cannot be directly applied for efficient bandwidth reservation. In a paper by Chao et al (1997), user mobility is estimated based on the aggregate history of handoff observed in each cell. The scheme assumes that each mobile will handoff to neighbouring cells with equal probability in the mobility-estimate-time-window and this assumption is not very accurate.

In our work, CAC and resource reservation schemes based on the probabilistic prediction of user mobility by using two different mobility models, parameters measurement model and fuzzy model are studied and their performance are analysed. In the mobility prediction approach used in this work, not only the cells where the mobile users will handoff but also when it will handoff are predicted.
2.1 CAC SCHEME IN WCDMA SYSTEM USING A MOBILITY MODEL BASED ON LZW ALGORITHM

2.1.1 Model Description

A mobile communication network with a cellular wireless infrastructure is considered. A handoff could fail due to insufficient bandwidth in the new cell, causing the handoff call to be dropped. The network topology, channel holding time distribution and user mobility pattern considered in our study are described in the following subsections. The capacity of a CDMA system is limited by the total interference it can tolerate, which is called the interference-limit system. Users with different traffic profiles and attributes, such as the service rate, the Signal-to-Interference ratio (SIR) requirement, media activity, etc., introduce a different amount of interference to the system. The Signal to Interference Guard Margin (SIRGM) is a natural extension of the guard channel idea developed in the context of TDMA/FDMA systems by considering the load factor for system capacity estimation in CDMA systems. The amount of SIRGM is dynamically adjusted by the RRE module.

i) Network Topology

In solving the CAC and bandwidth reservation problems, the cellular network is modeled as a connected graph \( G = (V, E) \), where the vertex-set \( V \) represents the set of base stations, each serving a single cell, and the edge-set \( E \) represents the adjacency between pairs of cells. Figure 2.1 shows an example network representation with vertex-set \( V = \{a, b, c, ..., n\} \) and edge-set \( E = \{(a, b), (a, c), ..., (n, l)\} \).
ii) Channel Holding Time

The channel holding time is defined as the time during which a new or handoff call occupies a channel in a given cell, and it is dependent on the mobility of the user. While this is similar to the call holding time in the fixed telephone network, it is often a fraction of the total call duration in a wireless cellular network and need not have the same statistical properties (Jedrzycki et al 1996, Langdon et al 1983). Most research work on CAC and bandwidth reservation assumes the channel holding time following an exponential distribution (Fang et al 1999, Viterbi 1995), which is independent and identically distributed for all cells. It is assumed that the channel holding time follows a general distribution, which allows the independent identically distributed exponential channel holding time assumption to be relaxed.

iii) User Mobility Pattern

The symmetric random walk model has been quite popular among researchers in characterizing individual movement behaviour (Papoulis 1991, Viterbi 1995). In such a model, a mobile user will move to any one of the neighbouring cells with equal probability after leaving a cell. This model does not take into account the trajectory and channel holding time of a mobile. In cellular mobile networks, the mobility of a user during a call can be
represented by a sequence of events, N, H_1, H_2, H_3, ... H_n, ... E, where N represents the event that a new call is admitted, H_n represents the event of a mobile’s n^{th} handoff and E represents the call termination event. Note that in some cases, there are no handoff events during the lifetime of a call and thus no H_n in the sequence of events. In this sequence, N = (m, i, t), where ‘m’ represents the mobile requesting the call, ‘i’ represents the original cell and ‘t’ represents the time when the call arrives; H_n = (T_k, j), where T_k is the relative time elapsed since the beginning of the call and ‘j’ is the cell to which the mobile will handoff; and E = (T_k),end of the call. We quantize the relative time into slots of equal duration ‘T’, a design parameter. So, T_k is the k^{th} time slot since the beginning of the call. In general, a mobile user usually travels with a specific destination in mind.

So, the mobile’s location and channel holding time in the future are likely to be correlated with its movement history. Therefore, in our model, the sequence of events N, H_1, H_2, H_3, ... H_n, ... E is assumed to be generated by a m^{th} order Markov source, in which the states correspond to the contexts of the previous ‘m’ events. The probabilities of possible next events can depend on a list of ‘m’ previous events.

2.1.2 Mobility prediction

In this part of our work, for predicting the mobility of the users, the LZW (Lempel-Ziv-Welch) algorithm for data compression is used. The equivalent character-based LZW algorithm builds in an online fashion a probabilistic model (or a tree) that feeds probability information to an arithmetic coder, which encode a sequence of probability of p using bits.

The sequence of events N, H_1, H_2, H_3, ... H_n, ... E during the lifetime of a call corresponds to the substring in the LZW algorithm. The mobility
database of every mobile at a specific time holds a Mobility Tree, which is a probability model corresponding to that of the LZW algorithm. Each node except for the root in the Mobility Tree preserves the relevant statistics that can be used to predict the probability of following events. As in data compression, the Mobility Tree of the mobile is built in an online fashion.

![Mobility Tree](image)

**Figure 2.2 A mobility tree used for mobility prediction**

When a mobile requests a new call, the predictor sets the current node to the root of the tree according to the mobile, cell and time, and calculates the probabilities of all possible events of this mobile. On seeing the actual event of the mobile, the predictor walks down the tree and is ready to predict again. When an event is not in the Mobility Tree, a prediction fault is generated and the tree is updated. Figure 2.2 shows an example Mobility Tree of mobile ‘m’ at cell ‘a’ in the time interval 8:00-9:00 a.m. When the mobile requests a new call in cell ‘a’ in the time interval 8:00-9:00 a.m., we can use the statistics preserved in the nodes of its Mobility Tree to predict the probabilities of the next possible events of this mobile: it will terminate the call without handoffs in the 2\textsuperscript{nd} time slot with probability of 2/56, handoff to cell ‘b’ in the 2\textsuperscript{nd} time slot with probability of 15/56, etc.
This mobility prediction scheme maintains the statistics in a tree. An important issue is how this model can be implemented. In fact, a tree is a multi-way tree with a path from the root to a unique node for each string represented in the tree. There are many ways to implement the nodes of a tree. The fastest approach is to create an array of pointers for each node in the tree with a pointer for each character of the input alphabet. This method can waste considerable memory space. An alternative is to use a linked list at each node, with one item for each possible branch. This uses memory economically, but can be more processing intensive. Some improvement may be achieved by moving an item to the front of the list each time it is used. In practice, in order to reduce the memory and computation complexity, it is desirable to limit the number of statistics used for mobility prediction. A sliding window may be used to ensure that only the most recent statistical data are involved in mobility prediction and the old data are discarded.

2.1.3 Call Admission Control and SIR Margin Reservation

2.1.3.1 Calculation of $P_{i,j}(T_k)$

The probability that a mobile originally in cell $i$ will visit cell $j$ during time slot $T_k$, i.e., $P_{i,j}(T_k)$ is calculated based on the predicted mobility of each user. From the Mobility Tree, we can see that a mobile taking different paths can visit certain cell in the same slot. Using total probability theorem (Ramjee et al 1996), we must add all of these probabilities to get $P_{i,j}(T_k)$. We show by an example how to get this probability.

**Example**: A mobile $m$ requests a new call at cell $a$ in the time interval 8:00-9:00 a.m. From the Mobility Tree in Figure 2.2, we can see that $m$ can take several different paths to visit cell $b$. We describe these paths by sequences of events:
Path 1: \( N(m, a, 8:00-9:00\ a.m.), H(T_1, b), E(T_2) \).

By Path 1, \( m \) will visit cell \( b \) in \( T_1 \) and \( T_2 \) with probability: \((3/56) \times (3/3) = 3/56\)

Path 2: \( N(m, a, 8:00-9:00\ a.m.), H(T_2, b), E(T_4) \).

By Path 2, \( m \) will visit cell \( b \) in \( T_2, T_3 \) and \( T_4 \) with probability: \((15/56) \times (5/15) = 5/56\)

Path 3: \( N(m, a, 8:00-9:00\ a.m.), H(T_2, b), H(T_2, d) \).

By Path 3, \( m \) will visit cell \( b \) in \( T_2 \) with probability: \((15/56) \times (6/15) = 6/56\)

Path 4: \( N(m, a, 8:00-9:00\ a.m.), H(T_2, b), H(T_5, d) \).

By Path 4, \( m \) will visit cell \( b \) in \( T_2, T_3, T_4 \) and \( T_5 \) with probability: \((15/56) \times (4/15) = 4/56\).

So, 
\[
P_{a,b}(T_1)=3/56
\]
\[
P_{a,b}(T_2)=(3/56)+(5/56)+(6/56)+(4/56)=18/56
\]
\[
P_{a,b}(T_3)=(5/56)+(4/56)=9/56
\]
\[
P_{a,b}(T_4)=(5/56)+(4/56)=9/56 \quad \text{and} \quad P_{a,b}(T_5)=4/56.
\]

When a mobile is active in cell ‘\( i \)’, we can get the **Most Likely Cell-Time** (MLCT) of that mobile, a cluster of time units at a cluster of cells when and where a mobile will most likely visit in the future. We select cells and time slots with \( P_{i,j}(T_k) \) greater than MLCT threshold , a design parameter, to form the MLCT of this mobile.

2.1.3.2 **CAC and SIR Reservation for New Calls**

When a new call arriving at mobile ‘\( m \)’ with a signal to interference ratio requirement ‘\( \gamma_n \)’ requires admission to cell ‘\( i \)’, the CAC algorithm first
checks if the signal to interference ratio margin \( \gamma_m \) (which is the difference between the current total average SIR of all the active calls and minimum required SIR for each active call) in the cell ‘i’ is greater than the required SIR \( \gamma_n \). The call is rejected if the cell does not have enough \( \gamma_m \). Otherwise, CAC will check the availability of \( \gamma_m \) in the MLCT of this mobile. Let \( \gamma_m (j, T_k) \) represent the SIR margin at cell ‘j’ in time slot ‘\( T_k \)’. The checking result can be written as:

\[
\text{Check}(j,T_k, \gamma_n) = \begin{cases} 
1 & \text{if } \gamma_m > \gamma_n \\
\gamma_m & \text{otherwise}
\end{cases}
\]  
\tag{2.1}

Based on these values, the new call will be admitted if the following holds:

\[
\sum_{i,k\in\text{MLCT}} P_{i,j}(T_k) \cdot \text{Check}(j, T_k, \gamma_n) \geq \alpha \sum_{i,k\in\text{MLCT}} P_{i,j}(T_k)
\]  
\tag{2.2}

where ‘\( \alpha \)’ is the admission threshold and should be controlled adaptively. When a new call is admitted, signal to interference ratio is reserved in the mobile’s MLCT. Let \( \gamma_r(j, T_k, m) \) denote the SIR reserved at cell ‘j’ in time slot \( T_k \) for mobile ‘m’.

The reserved SIR is:

\[
\gamma_r(j, T_k, m) = P_{i,j}(T_k) \cdot \gamma_n + 0.2 \cdot \gamma_n
\]  
\tag{2.3}

The second term in the above expression indicates the amount of resource reserved for the traffic originated in the neighbouring cells. The availability of SIR in MLCT is updated accordingly.
2.1.3.3 Adaptive Control of Admission Threshold

The mobility prediction functions may not work well for some mobile users, especially those who do not have favourite routes. Moreover, if ‘α’ is too small, the handoff dropping probability may exceed the target value; if ‘α’ is too large, resource utilization will be decreased. So, admission threshold should be controlled adaptively. The handoff dropping probability of mobile \( m \), \( P_{hd}(m) \) is calculated by dividing the number of handoff drops to the total number of its calls in the recent history. Let \( P_{hd\_target}(m) \) denote the target value of handoff dropping probability of mobile \( m \). If \( P_{hd}(m) < P_{hd\_target}(m) \), admission threshold is decreased by \( \varepsilon \), a design parameter; otherwise, it is increased by \( \varepsilon \). By adaptive control of ‘α’, a better balance of guaranteeing \( P_{hd} \) and maximizing resource utilization is achieved.

When a mobile \( m \), with SIR requirement \( \gamma_{ho} \), requires handoff to cell \( i \), the CAC algorithm will admit it if the \( \gamma_m \) of cell \( i \) is greater than \( \gamma_{ho} \). Then, CAC will calculate \( P_{i,j}(T_k) \) and get the MLCT of \( m \) based on the Mobility Tree. SIR margin is reserved for \( m \) in its MLCT accordingly.

2.2 CAC SCHEME IN WCDMA SYSTEM USING MOBILITY MODEL BASED ON THE DIRECTION AND SPEED OF THE MOBILE

The mobility model used in this part of our work finds the candidate cell by using the speed and direction of movement of the mobile users. The probability of the mobile user moving from one cell to the other cell in the specific time slot is estimated by this mobility model. The probability values obtained from this model are used by the call admission control and resource reservation system to improve its performance.
2.2.1 Determination of Reservation Area

In order to avoid premature termination of the call without the intention of the users, channels must be reserved in the surrounding area. This part of the terrain is called the reservation area for the user. The construction of the reservation area is a two step process: finding the candidate cell in which a reservation is advised (but not required), and the calculation of the likeliness value for each reservation request.

2.2.2 Finding the candidate cells

Finding the set of candidate cells into which a user may move depends on observing the following important points.

- A user tends to keep his direction since he moves towards a destination, though there will be small changes in his direction due to obstacles in the terrain.

- The obstacle in the terrain affect all individuals of a society in the same manner. i.e., if a high way has an arc to the left, all the vehicles draw an arc to the left.

- The frequency of change in the direction signifies the type of structure the user is on. A user on the highway follows a path close to a straight line where a user on the streets often changes his direction.

The reservation area for a user is an elliptical region surrounding the mobile station as shown in the Figure 2.3. An ellipse can be defined by its size in both major and minor axes, the orientation and coordinates of one of its two foci. The orientation is the angle between the major axis of the ellipse
and the x-axis of the coordinate system. The size in the axes and orientation of the ellipse are determined by the mobility pattern of the user. Of the two foci of the ellipse, the first one is located at the current locus of the user and the ellipse is drawn such that the second focus is on the moving direction of the user. In other words, the direction of the user determines the orientation of the elliptical region. The length in the major axis can be determined from the speed of the user and the length of the duration for which the system tries to make reservation. Higher speeds and longer reservation duration tend to increase the length in the major axis. The width in the minor axis can be determined from the length in the major axis and the mobility pattern of the user in the near past. Thus, the width of the ellipse depends implicitly on the speed and reservation duration in addition to the mobility pattern of the user. The reservation area of the user who has been keeping its direction for some time will be a very narrow ellipse while the reservation area of a subscriber who has been changing its direction frequently will be wider. In the extreme case, the reservation area will be a circle. Once the reservation area has been determined, the set of candidate cells can be determined by finding the cells intersected by the reservation area. In a hexagonal grid, this task can be accomplished by checking whether the center of the cell is inside a hypothetical ellipse, co-centered with the reservation area, but larger by 2R in length and width, where ‘R’ is the cell radius.

![Figure 2.3 Mobility model with Reservation Area](image-url)
2.2.3 Calculation of the Likeliness Value

The likeliness that a subscriber will be at a specific point is related to the distance and the angle between the point and current position of the user. The distance and likeliness value are inversely proportional, since the probability that the user changes his direction increases with distance. A point on the moving direction of the user is more likely to be visited with respect to a point on the side. Therefore, the angle between the moving direction of the user and the line connecting the current location of the mobile and the point of concern is inversely proportional with the likeliness. The likeliness value can be calculated as:

\[
P_{i,j}(T_k) = w_1 \cdot \left[1 - \frac{d_j}{(2a + R)}\right] + w_2 \cdot \left[1 - \frac{\theta_1 - \theta_2}{\pi}\right]
\]

(2.4)

where

- \(P_{i,j}(T_k)\) is the likeliness (probability) that the mobile in the cell ‘i’ visit the cell in the time slot ‘\(T_k\)’.
- \(d_j\) is the distance between the mobile and cell ‘j’.
- \(R\) is the radius of the cell.
- \(w_1\) is the weight associated with the mobile-cell center distance.
- \(w_2\) is the weight associated with the location of the cell center with respect to the direction of the mobile. (Both \(w_1\) and \(w_2\) are equal to \(\frac{1}{2}\)).
- \(\theta_1\) is the angle between the line that connects the mobile to the cell center and x-axis.
- \(\theta_2\) is the direction of the mobile and |\(\theta_1 - \theta_2\)| < \(\pi\).
- \(a\) is the semi-major axis (half of the length) of the ellipse representing the reservation area.

The calculation of likeliness value given in equation(2.4) is explained in Appendix-I.
When a mobile is active in cell ‘i’, we can get the Most Likely Cell-Time (MLCT) of that mobile. We select cells and time slots with $P_{i,j}(T_k)$ greater than MLCT threshold, a design parameter, to form the MLCT of this mobile.

### 2.2.4 Capacity and Load Estimation in CDMA Systems

The capacity of a CDMA system is limited by the total interference it can tolerate. Users with different traffic profiles and attributes, such as the service rate, the Signal-to-Interference ratio (SIR) requirement, media activity, etc., introduce different amount of interference to the system. The effect of interference increase for traffics with the following attributes for user ‘i’ is studied in literature.

- $R_i$: service (data) rate for user $i$.
- $\varepsilon_i$: target signal-to-interference ratio (SIR).
- $\nu$: media activity level
- $f$: other cell to own cell interference ratio

To derive the amount of interference $\Delta I_i$ introduced by user ‘i’ with data rate $R_i$ and target SIR, we have

$$\varepsilon_i = (E_b / N_0)_i = \frac{S_i/(\nu_i R_i)}{(I_{\text{total}} - S_i)/W}$$  \hspace{1cm} (2.5)

where $I_{\text{total}}$ is the total received power from $N$ active users at the base station, i.e.

$$I_{\text{total}} = \sum_{i=1}^{N} S_i + P_N$$  \hspace{1cm} (2.6)

$E_b$ is the energy per user bit, $N_0$ the noise spectral density, $S_i$ and $\nu_i$ are the received power at the base station from user $i$ and the activity level of user $i$, respectively.
respectively, and \( P_N \) is background noise. \( I_{\text{total}} \) is limited by an upper-bound \( I_{\text{Th}} \) for a system. When \( I_{\text{total}} \) is higher than the upper-bound, the system is unstable and the overall interference increases dramatically. By rewriting the equation 2.6, we can express the received power \( S_i \) for user \( i \) at the base station as

\[
S_i = (1 + \frac{W}{\varepsilon_i, \nu_i, R_i})^{-1} I_{\text{total}} = \Delta \rho_i I_{\text{total}} \tag{2.7}
\]

where

\[
\Delta \rho_i = (1 + \frac{W}{\varepsilon_i, \nu_i, R_i})^{-1} \tag{2.8}
\]

is called the load factor increment (Holma and Toskala 2000). The current load factor of such an interference system is the sum of load factor increments brought by \( N \) active mobile users,

\[
\rho = \sum_{i=1}^{N} \Delta \rho_i \tag{2.9}
\]

(Shapira and Padovani, 1994) and (Holma et al 2002) estimated the interference increase by taking into account the load curve as shown in Figure 2.4. The ratio of \( I_{\text{total}} \) to background noise \( P_N \) is called noise-rise and denoted by \( \eta \). The maximum value of \( \eta \), denoted by \( \eta_{\text{max}} \), is normally set to 10 (Holma and Toskala 2000). The noise-rise \( \eta \) is

\[
\eta = \frac{I_{\text{total}} + P_N}{P_N} = \frac{\sum_{i=1}^{N} S_i + P_N}{P_N} = (1 - \rho)^{-1} = (1 - \sum_{i=1}^{N} \Delta \rho_i)^{-1} \tag{2.10}
\]

By taking the partial derivative of \( I_{\text{total}} \) with respect to \( \rho_i \), we get
The interference increment, $\Delta I_i$, can be expressed in terms of $R_i$ as shown below:

$$
\Delta I_i = \Delta \rho_i I_{total} = \frac{\partial I_{total}}{\partial \rho_i} = \frac{\partial (P_N/(1-\rho))}{\partial \rho_i} = \frac{P_N}{(1-\rho)^2} = \frac{I_{total}}{1-\rho}.
$$

(2.11)

The load curve serves as a good tool for interference increment estimation in our proposed scheme. The key point here is that given a data rate $R_i$, the load increment $\Delta \rho_i$ can be computed to yield interference increment $\Delta I_i$.

$$
\Delta I_i = \frac{\Delta \rho_i I_{total}}{1-\rho} = \frac{(1 + \frac{W}{\varepsilon V_i R_i})^{-1}}{1-\rho} I_{total}
$$

(2.12)

2.2.5 Preferential Treatment with IGM

In this section, an efficient radio resource management scheme based on the concept of interference guard margin (IGM) is discussed. In a mobile communication system with $N$ active mobile users, the $i^{th}$ ($i < N$) user’s traffic profile, which characterizes its services, is described as

$$
S(i) = \{r_i, \{R_{max}, R_{min}\}_i, I_{I_i}, M_i\},
$$

(2.13)
where $r_i$, $(R_{\text{max}}, R_{\text{min}})_i$, $\Pi_i$ and $M_i$ in $S(i)$, denote user $i$'s rate adaptivity, service rate range, priority and mobility, respectively. The proposed service model is designed to take advantage of modern coding schemes and advanced mobile communication technologies as described below.

First, $r_i$ is a binary indicator that indicates whether the user can be serviced at reduced bit-rates when the system is congested. Second, the service rate range $(R_{\text{max}}, R_{\text{min}})_i$ describes the target bandwidth consumption. If the network has enough resources, the request can be admitted at $R_{\text{max}}$. If the cellular system is overloaded (congested), a rate-adaptive user can be serviced at a lower rate (down to $R_{\text{max}}/2$ or even $R_{\text{min}}$) with degraded quality of service. Adaptation only takes place at the time of admitting new calls or at handoff events. Third, the priority tag $\Pi_i$ helps the system to identify high priority users, who are likely to receive better QoS guarantees. Each different mobility traffic has a different weighting factor to estimate the amount of resources necessary to be reserved.

2.2.5.1 Interference Guard Margin (IGM)

IGM is a natural extension of the guard channel idea developed in the context of TDMA/FDMA systems by considering the load factor for system capacity estimation in CDMA systems. As illustrated in Figure 2.4, we have the following two operations. First, the load curve is used to estimate the load increase as well as the interference increase. Second, a certain amount of IGM, instead of guard channels, is pre-reserved for high priority calls. The amount of IGM is dynamically adjusted by the RRE module.

For a new call to be admitted, the total interference level should not exceed the upper bound of the interference with threshold $I_{\text{Th}}$ that the system can tolerate. In addition to the constraint of $I_{\text{Th}}$, a lower priority call should
comply with the augmented constraint $I'_{Th}$. The margin between $I_{Th}$ and $I'_{Th}$ is exactly the guard margin, which provides the preferential treatment to high priority calls by limiting the access to the low priority calls. The interference guard margin is given as,

$$IGM(j) = \alpha \sum_{i \in MLCT} P_{i,j}(T_k) \Delta I_{min,i}$$

(2.14)

### 2.2.5.2 Call Admission Control Algorithm

The CAC algorithm for a media type with three different data rates is discussed in this topic. In this, a new call or a handoff call can be admitted into the system with three data rates: $R_{max}$, $R_{half}$ and $R_{min}$. It can be generalized to a media type consisting of even more rates. Note that $IGM_{new}$ and $IGM_{handoff}$ are the estimated interference margin required to be reserved for new and handoff calls, respectively. The basic concept behind CAC is to test whether there is enough system resource left to serve the current call request at a certain rate, after reserving the necessary resource for preferential treatment. The CAC test is performed according to the following steps:

**Step 1. New or handoff call test:** An incoming call is first identified as a new or a handoff call type to decide its priority.

**Step 2. Rate adaptivity test:** The rate adaptivity of a new call (handoff call) is tested to decide whether it can be serviced at a lower data rate if the system is congested.

**Step 3. Non rate-adaptive call test:** If the call is rate-adaptive, go to Step 4. Otherwise, test whether the amount of interference after
admitting the current call and reserving the estimated IGM, will exceed the maximum interference level that the system can tolerate.

**Step 4. Rate adaptive call test**: If the call is rate-adaptive, the current call could be serviced at rates of \( R_{\text{max},i} \), \( R_{\text{half},i} \) and \( R_{\text{min},i} \), depending on the system traffic condition. The amount of interferences introduced by a call are \( \Delta I_{\text{max},i} \), \( \Delta I_{\text{half},i} \) and \( \Delta I_{\text{min},i} \) when it is serviced at rates \( R_{\text{max},i} \), \( R_{\text{half},i} \) and \( R_{\text{min},i} \) respectively. Then, the admission criteria by the order of data rates for the highest to the lowest is tested. The call is served at its highest admissible rate.

### 2.3 PARAMETERS MEASUREMENT BASED CAC SCHEMES IN WCDMA SYSTEM

In the broadband wireless networks, providing multimedia services with a quality of service (QoS) guarantee in wireless environment is more challenging due to the limitation in the bandwidth, mobility of the users and dynamic characteristics of the channel. In the literature, QoS in wireless networks is considered at two levels, i.e. the application level and connection level. Application level QoS is related to perceived quality at the user end and is commonly considered in packet switched networks. The parameters, such as delay/delay jitter, error/loss and throughput, etc., are used to describe the application level QoS. Efficient packet access protocols and packet scheduling schemes play key roles in solving these QoS problems. Connection level QoS is related to connection establishment and management. It measures the connectivity and continuity of service in a wireless network, and it is defined by two parameters: the new call blocking probability, which measures the service connectivity and the handoff dropping probability, which measures the service continuity during handoff.
In this part of the work, a system which provides connection level QoS according to service requests from the users, under the constraint of limited and varying bandwidth resources is analysed. The main features of the system are highlighted as follows.

- It is based on a service model consisting of three service classes i.e., class-1 (real-time-high priority), class-2 (real-time-low priority) and class-3 (best effort-non real time).
- It deploys different resource-reservation schemes adaptively for real-time service classes (i.e., class-1 and class-2) to guarantee their connection-level QoS.
- It uses an efficient dynamic call admission control scheme to meet the target handoff-dropping probability for real-time services.
- It exploits the rate-adaptive feature of multimedia applications to further improve the efficiency of resource utilization.

2.3.1 Dynamic QoS Management System

The block diagram of the QoS management system considered in our work is illustrated in Figure 2.5. With QoS as the kernel, this system allows different applications to request different QoS from the network through a service model. Application profiles are mapped into the service model by different forms of traffic specifications. Network resources are adaptively allocated to different service classes by employing adaptive resource-allocation schemes, including call-admission control and resource reservation, according to the service model and QoS requirements. The adaptation module enables the negotiation of QoS between applications and
networks whenever it is necessary. Each component is described in detail below.

Figure 2.5 Block diagram of the Dynamic QoS management system

The service model used in this work for multimedia applications in wireless networks is described below. First, multimedia applications are classified into real-time and non real-time applications according to their delay requirements. In order to achieve desired QoS for a real-time application, it is usually necessary to maintain a minimum data rate requirement during its lifetime. The setup of a connection requires call-admission control and resource reservation to prevent network congestion and dropping of ongoing calls. The handoff-dropping probability is considered as the primary QoS requirement and it is assumed that, it has more significant impact on the overall connection-level QoS measurement. For non real-time applications, we use the best-effort service adopted in traditional IP networks. To improve resource utilization of the entire network, the resources reserved for real-time applications but not yet being in use is utilized to carry non real-time data. The different applications are categorized into the following three classes based on our earlier discussion.
• Class-1 traffic: represents real-time applications that require absolute continuity, i.e., handoff dropping probability is to be zero.

• Class-2 traffic: represents real-time applications that can tolerate a reasonably low handoff-dropping probability.

• Class-3 traffic (best effort traffic): represents non-real-time applications that do not need a minimum bandwidth to set up a connection.

2.3.1.1 Application Profile

The above service model covers many application-level QoS aspects, such as delay, priority, and bandwidth adaptation, as well as pricing and mobility aspects. The network uses different application profiles for different service classes. For real-time service classes, including both class-1 and class-2, the minimum required bandwidth to meet the delay requirement is necessary. The application profile also includes the required handoff-dropping probability for real-time service classes. For the class-1, the target handoff-dropping probability should be zero. For the class-2, the target handoff-dropping probability is bounded by $P_{\text{hd, target}}$. Moreover, an application requesting the class-1 service should also provide its mobility information, so that the network could predict the cells that the mobile is going to visit during its lifetime. For the best-effort service class, there is no minimum bandwidth requirement. In this class, the traffic load is described by the packet-generation rate and the packet size.
2.3.1.2 Resource Allocation

For real-time service classes, resource allocation includes call-admission control (CAC) and resource-reservation (RR) mechanisms. These two mechanisms are closely related to each other to achieve the desired QoS for a given application. A different resource-allocation scheme, as explained below, is used for each service class to provide appropriate QoS to the corresponding applications.

- For the class-1 traffic, it is necessary to reserve resources in other cells the mobile host may visit, which is indicated by its “application profile”. The reserved resources can only be used by the class-1 traffic or class-3 (best-effort) data call until the reserving class-1 call arrives. This guarantees resources to each class-1 call upon handoff. CAC is simply based on whether resources are reserved successfully.

- For the class-2 traffic, aggregate resources are reserved for the handoff calls of this class to maintain a reasonably low target handoff-dropping rate. A measurement-based algorithm that dynamically adjusts the threshold in the reservation scheme is described in the next section.

- The class-3 (best-effort service) traffic is serviced with the remaining resources, including those reserved but not being used by real-time service classes. The above two real-time service classes can pre-empt this service class, thus improving the overall utilization of the network without sacrificing the QoS guarantee for real-time traffic.
In this scheme, there are four types of real-time traffic listed in the decreasing order of priorities: the class-1 handoff call (1-H), the class-2 handoff call (2-H), the class-2 new call (2-N), and the class-1 new call (1-N). The resource sharing among the real-time traffic is performed by our resource-allocation scheme, as follows. The reservation for 1-H calls is similar to the complete partitioning (CP) policy (Epstein and Schwartz 1995). It ensures enough resource for admitted 1-H calls. For 2-H calls, a post reservation (Epstein and Schwartz 1995) for a lower handoff-dropping probability requirement is achieved by the proposed dynamic resource reservation scheme. The 2-N and 1-N calls adopt the complete sharing (CS) (Epstein and Schwartz 1995) policy to use the remaining available resources.

2.3.1.3 Rate-Adaptive Applications

The wireless network is a highly variable environment where available link bandwidth may vary with network load and channel condition. By using rate-adaptive features of many multimedia applications, the resource-allocation scheme considered in this work, can be adaptive to network conditions. In this system, when rate-adaptive applications make a connection request to the network, they specify the minimum and maximum data rate required to be supported by the network. Adaptation first takes place while admitting a new call. If the network has enough resources available, the request is admitted. Otherwise, it is admitted with a lower data rate. If the network is overloaded and cannot be satisfied, the call is blocked. A rate-adaptive connection admitted at could be handed off at a lower rate if the cell it is entering is heavily loaded. On the other hand, a call admitted at could be upgraded to a higher rate if the cell it is going to enter is underutilized. The rate adaptation for a call is, only upon its admission and handoff.
2.3.2 Parameters Measurement Based Dynamic Interference Guard Margin Scheme

In this section, a parameters measurement based dynamic interference guard margin scheme for the WCDMA system, which is designed for the class-2 traffic to meet its target handoff-dropping probability is described.

2.3.2.1 Dynamic Interference Margin Scheme

It has been justified in the literature that the guard channel scheme plays an important role in the new call-blocking probability ($P_{nb}$) and handoff-dropping probability ($P_{hd}$). For a single service-class network, increasing the admission threshold ($T_h$) results in less resources reserved for handoff calls, thus strictly increasing the $P_{hd}$. At the same time, the $P_{nb}$ decreases accordingly, because more resources become available for new arriving calls. On the other hand, decreasing the threshold has the opposite effect. Moreover, the overall system utilization could also be influenced by the value of threshold. Reserving more resources than needed by handoff calls results in lower system utilization since reserved resources cannot be used by new call requests. Thus, the selection of $T_h$ is critical to the system design. The value of $T_h$ should be selected so that necessary and sufficient resources are reserved for handoff calls to meet their QoS requirements in terms of $P_{hd}$. Apparently, the selection of $T_h$, i.e. the shared reservation of resources for handoff calls, should dynamically vary with changing traffic conditions. In CDMA system, the reservation is to be made in terms of interference margin. The load curve given in the previous section 2.2.4 in the Figure 2.4 is a good tool for estimating the interference in this scheme also. The key point here is that given a data rate $R_i$, the load increment $\Delta \rho_i$ can be computed to yield interference increment $\Delta I_i$. 
For a new call to be admitted, the total interference level should not exceed the upper bound of the interference with threshold $I_{Th}$ that the system can tolerate. In addition to the constraint of $I_{Th}$, a lower priority call should comply with the augmented constraint $I'_{Th}$. The margin between $I_{Th}$ and $I'_{Th}$ is exactly the guard margin, which provides the preferential treatment to high priority calls by limiting the access to the low priority calls.

The above discussion is based on the assumption of a single traffic class. Here it is extended to multiple traffic classes by applying a single threshold to reserve resources for class-2 handoff calls in a post reservation fashion. The remaining resources are used by the new call requests from both class-1 and class-2 calls in a complete sharing fashion. Decreasing the value of $I'_{th}$ in our multiple traffic class network will reserve more resources for class-2 handoff calls. Unlike the single traffic network considered in Oh and Tcha (1992), this will intuitively result in the lower or the same $P_{hd}$ for class-2 handoff calls and higher or the same $P_{nb}$ for class-1 new call and class-2 new calls. Similarly, increasing has the opposite effect. Thus, the objective of our dynamic interference margin scheme can be stated as follows: For a given system with a target handoff-dropping probability for class-2 handoff calls, threshold should be selected to keep the resulting $P_{hd}$ for class-2 handoff calls as close to $P_{hd, target}$ as possible without exceeding it, while the $P_{nb}$ for new class-1 and class-2 calls should be minimized.

### 2.3.2.2 Measurement of Handoff-Dropping Probability

There are two approaches used to estimate the handoff-dropping probabilities: modelling based and measurement based. The modelling-based approach (Kwon and Choi 1998, Lin et al 1995) uses theoretical models to deduct the probability via mathematical analysis by assuming some parameters of the model. The measurement-based approach uses observed
network conditions, obtained by some measurements, to do simple estimation. The performance of this approach is usually good considering its less computational complexity and its adaptability to changing conditions of practical networks. For these reasons, a measurement-based algorithm that aims at achieving the objective defined above by dynamically adjusting the threshold is considered. Basically, the call-admission control scheme discussed in this section is based on measurements of the current value of the dropping probability $P_{hd}$. Compared with other measurement-based schemes which rely on the measurement of current traffic conditions (Naghshineh and Schwartz 1996, Kwon et al 1999), measuring $P_{hd}$ directly gives more accurate information and enables more efficient control in order to achieve our objective.

At the beginning, an initial value of $I_{th(initial)}$ is selected for a given cell. The BS of this cell monitors its $P_{hd}$ for the class-2 hand off calls. When the measured $P_{hd}$ reaches or exceeds the target value $P_{hd, target}$, $I_{th}$ is decreased by one unit (in dB) so that more amount of resource can be reserved for handoff requests. Otherwise, $I_{th}$ is increased by one unit (in dB) to admit more new call requests. The proposed algorithm is summarized with the pseudocodes given in Figure 2.6, where $I_{total}$ is the current total interference observed at the base station and $I_{new}$ the amount of interference which will be added due to the acceptance of a new call. Some design issues are discussed in the following subsections.

i) **Update Frequency**

An important design issue that affects the performance of the scheme is how often and when to update the measurement of $P_{hd}$ to dynamically adjust $I_{th}$. A frequent update can keep pace with changing traffic conditions. If the $I_{th}$ is not updated quickly enough to match the variation in
system conditions, it could result in lower channel utilization (due to late increase in) or higher (due to late decrease in). However, a new value of $I_{th}$ does not immediately affect the measured value of $P_{hd}$. Thus, very frequent updates could cause unnecessary fluctuations and burden the system. Therefore, a trade-off between fast response and system stability is desired.

Whenever a handoff request is dropped, the BS checks the current value of $P_{hd}$ and decreases $I_{th}$ if $P_{hd} \geq P_{hd\_target}$. Since only handoff-dropping events could drive the $P_{hd}$ higher than the target value, the system takes the corrective action by decreasing the threshold to avoid further handoff drops. At the same time, a timer is set. If there are no further handoff drops upon the expiration of the timer, the BS checks if $P_{hd} < P_{hd\_target}$. If it is true, the threshold is increased to improve channel utilization. Otherwise, it renews the timer. Before the timer’s expiration, if there are further handoff drops, the BS again checks $P_{hd}$. If necessary, it further decreases $I_{th}$ and resets the timer.

ii) Timer Setting

The timer setting allows the effect of decreasing being reflected in the change of $P_{hd}$ since it may take several successful handoff calls before the measured $P_{hd}$ falls below the target $P_{hd\_target}$ or before it gets to the steady state. The timer is renewed when the $P_{hd}$ continues to drop (indicated by no handoff dropping) until it falls below $P_{hd\_target}$, which triggers an increase in $I_{th}$, or further handoff dropping happens when $P_{hd} \geq P_{hd\_target}$, which means reserved resources are not enough for handoff and a further decrease in $I_{th}$ is needed. Thus, it avoids the unnecessary fluctuation of $I_{th}$ and provides necessary updating. The value of the timer is the maximum time that the system could be underutilized.
iii) Initial Threshold

Another issue is the selection of the initial threshold $I_{th}$. Theoretically, it can be derived from mathematical analysis conducted in the previous section according to given $P_{hd\_target}$. This approach can quickly lead the system into a steady state. However, it requires modelling parameters such as the new call and the handoff-arrival rates, the cell-residence time, etc. Practically, when these parameters are not available, we can simply set $I_{th\_initial}$ to the $I_{max}$. Starting with $I_{th\_initial} = I_{max}$ maximizes system utilization from the beginning. However, it might be achieved at the cost of violating $P_{hd\_target}$ the bound in the initial period.

Note that the CAC scheme, as illustrated in Figure 2.6 is applied to the class-2 service to achieve its target handoff-dropping probability in our system. However, the new call admission of the class-1 service is also subjected to the dynamic changing threshold derived here.

2.4 FUZZY MODEL BASED CAC SCHEMES IN WCDMA SYSTEM WITH ADAPTIVE MULTI CLASS TRAFFIC

Call Admission Control in CDMA cellular networks has been specifically addressed as a means of managing interference (load control) to ensure efficient QoS support, as the capacity of CDMA systems is interference-limited. CAC directly controls the number of users in a cell in order to keep the interference under a tolerable limit so that an adequate radio link performance and required QoS for each user can be maintained. This clearly indicates that there is a trade-off between the system capacity and the overall communications quality, that is, Grade of service (GoS) and QoS in effect. In general, due to the time-varying soft capacity of CDMA systems, the decision of CAC exhibits an error in both directions: to accept and to reject.
01 Initialize $I_{th} = I_{max}$
02 SET TIMER
03 WAIT FOR CALL REQUEST ARRIVAL
04 If NEW CALL REQUEST ARRIVES
05 CHECK THE RATE ADAPTIVENESS
06 If IT IS NON RATE ADAPTIVE CALL
07 ESTIMATE $I_{new}$ AND CHECK $(I_{total} + I_{new}) \leq I_{th}$
08 If YES ACCEPT THE CALL
09 else REJECT THE CALL
10 else ESTIMATE $I_{new}$ WITH $R_{max}$ AND CHECK $(I_{total} + I_{new}) \leq I_{th}$
11 If YES ACCEPT THE CALL
12 else ESTIMATE $I_{new}$ WITH $R_{min}$ AND CHECK $(I_{total} + I_{new}) \leq I_{th}$
13 If YES ACCEPT THE CALL
14 else REJECT THE CALL
15 If HANDOFF CALL REQUEST ARRIVES
16 CHECK THE RATE ADAPTIVENESS
17 If IT IS NON RATE ADAPTIVE CALL
18 ESTIMATE $I_{handoff}$ AND CHECK $(I_{total} + I_{handoff}) \leq I_{max}$
19 If YES ACCEPT THE HANDOFF CALL
20 else REJECT THE HANDOFF CALL
21 else ESTIMATE $I_{handoff}$ WITH $R_{max}$ AND CHECK $(I_{total} + I_{handoff}) \leq I_{max}$
22 If YES ACCEPT THE HANDOFF CALL
23 else ESTIMATE $I_{handoff}$ WITH $R_{min}$ AND CHECK $(I_{total} + I_{handoff}) \leq I_{max}$
24 If YES ACCEPT THE HANDOFF CALL
25 else REJECT THE HANDOFF CALL
26 UPDATE DROPPING PROBABILITY
27 If $P_{hd} \geq P_{hd\_target}$, DECREASE $I_{th}$ AND
28 RESET THE TIMER AND GO BACK TO 03
29 else CHECK THE TIMER
30 if TIMER EXPIRES
31 INCREASE $I_{th}$ AND RESET THE TIMER AND GO BACK TO 03
32 Else GO BACK TO 03

Figure 2.6 CAC Algorithm with Dynamic Interference Margin
Thus CAC limits and regulates the number of users admitted into the network under the constraints of network capacity and resource allocation strategies. Efficient CAC schemes can enhance the system capacity and service quality cost-effectively. In advanced multimedia cellular networks, the success of service deployment also depends on network capabilities in providing differentiated QoS to meet a wide range of customer demands. CAC schemes are vital for supporting the two-fold objective of cellular networks operation: (i) to deliver the required QoS; (ii) to provide operators freedom in controlling the payload and group behaviour of traffic classes regarding the required services as well as subscriber classes for optimizing network utility and revenues. In the wireless system, the decision made by the call admission controller is to be based on the imprecision and uncertain measurements due to user mobility, dynamic QoS requirements and varying channel conditions. Having the nature of coping with uncertainty and imprecision problems, fuzzy logic is expected to provide a good solution to the development of a call admission control scheme.

2.4.1 System Model

A wideband CDMA cellular system with hexagonal cells of equal size is considered. Each cell contains a centrally located base station (BS) with an omni directional antenna and the same radio spectrum is reused in all the cells. The mobile stations (MS) communicate with their home BSs via the air interface, and all the BSs are connected to a mobile switching center (MSC), which is in turn connected to backbone wire line network. Separate frequency bands are used for the forward and reverse links, so that each BS experiences interference only from the MSs. The focus is placed on the CAC process for the reverse link. In the forward link, each BS broadcasts a unique pilot signal to the MSs. The pilot signals from different BSs are distinguished by different scrambling codes. The MS can detect the pilot signal from any
BS when the strength of the pilot signal is above a certain level. Prior to transmission, an MS monitor the received pilot signal power levels from nearby BSs, and chooses its home BS according to the maximum received pilot signal power. It is assumed that MSs, BSs, and MSCs are properly designed such that, while tracking the signal from the home BS, an MS searches for all the possible pilots and maintains a list of all pilots whose signal power levels are above a prescribed threshold. This list is transmitted to the MSC periodically through the home BS. The MSC uses this information to make a decision on when handoff should start and also to estimate the mobility of the users.

When a call request arrives, the MSC uses the mobility information of all the active mobile users in the cell and also in the neighbouring cells together with their traffic characteristics, to calculate the resources utilized by all the active calls. It estimates the resource requirement of the new call and also makes reservation for the handoff calls. Then it finds the resource availability to accommodate the new call without degrading QoS of the calls already in service. For simplicity of presentation, we will focus on the mobility information and the interference introduced by the MS at its home BS (denoted as BS0) and at the six neighbouring BSs in the first-tier (denoted as BS1, BS2, ... BS6). Let $d_l(t)$, $l = 0, 1, ... 6$, denote the distance between the MS and BS$l$ at time ‘t’. It is assumed that, with a properly designed transceiver, the channel disturbance is mainly due to shadowing and path loss.

The local mean of the pilot signal power from BS1 received at the MS can be expressed as (Rappaport 2002)

$$a_1(t) = \gamma \left[ \frac{d_1(t)}{D_0} \right]^{\gamma} \cdot 10^{\frac{e(t)}{10}} + \nu_1(t)$$  \hspace{1cm} (2.15)
where $\gamma_l$ is a constant proportional to the transmitted signal power, $r$ is the path loss exponent; $D_0$ is the close-in reference distance that is determined from measurements close to the transmitter, and $\varepsilon_l(t)$ in dB at any ‘t’ is a Gaussian random variable (with zero mean and standard deviation $\sigma_s$) characterizing the shadowing phenomenon. For $l \neq k$, $\varepsilon_l(t)$ and $\varepsilon_k(t)$ are independent random processes. If the transmit power levels of all the pilot signals are the same, then $\gamma_l = \gamma$ for $l = 0, 1 \ldots 6$. $\nu_l(t)$ represents background noise power and multiple access interference (MAI) from the information-bearing signals in the forward link to all the MSs. When there are a large number of users in the system, the MAI can be modelled approximately by a Gaussian random process. Similarly, in the reverse link, the propagation loss $\Lambda_l(t)$ from the MS to BS $l$ is proportional to the product of the $r$th power of distance and a log-normal component characterizing the shadowing phenomenon, given by

$$\Lambda_l(t) = d_{l,t}(t).10^{\frac{\zeta_l(t)}{10}}$$  \hfill (2.16)

where $\zeta_l(t)$ in dB at any $t$ is a Gaussian random variable with zero mean and standard deviation $\sigma_s$. For the MS at a point M, suppose at time $t$, the received power of the signal from this MS at its home base station BS$_0$ is $P_{r,0}(t)$, then the received power of the signal from this MS at the neighbouring BSs can be expressed as

$$P_{r,l}(t) = P_{r,0}(t) \left[ \frac{d_l(t)}{d_0(t)} \right]^r 10^{\frac{(\varepsilon_l(t)-\varepsilon_0(t))}{10}}$$  \hfill (2.17)

where $l = 1, 2, \ldots, 6$

In the wideband CDMA cellular system, soft handoff is adopted since it can extend CDMA cell coverage and increase reverse link capacity. Soft handoff happens when an MS moves into the overlapping cell coverage area of two or
more BSs. The pilot channel’s bit energy to noise-plus-interference density ratio \((E_b/N_0)_p\) is used as the handoff measurement quantity. For simplicity, it is assumed that, in soft handoff, an MS is connected to two nearest BSs with strongest pilot signals, while it is power controlled by the BS that requires it to transmit at the smaller power. Consider that an MS moves into the soft handoff region from cell ‘l’ to cell ‘k’. If cell ‘k’ does not have enough resources to accept the handoff call when the MS moves out from the overlapping region, the call has to be dropped to guarantee the QoS of the existing calls in cell k. It should be mentioned that, during the soft handoff, the signal transmitted by the MS is received at both the BSs. With selection diversity at the receiving end, the required transmitted power of the MS (and, hence, the system resource) is reduced as compared with that in the hard handoff situation.

The SIR of each and every connection depends on the power emitted by mobile users, inter-cell and intra-cell interferences and the thermal noise. The transmit power of the mobile station can be calculated from the SIR requirement of the user, which is given by,

\[
\frac{E_b}{N_0} = \frac{W_c \cdot P_j}{\nu_j R_j \left( I_{Total} - P_j \right)}
\]

where \(W_c\) is the chip rate, \(P_j\) is the received signal power from user \(j\), \(\nu_j\) is the activity factor of user \(j\), \(R_j\) is the bit rate of user \(j\) and \(I_{Total}\) is the total wideband power received including thermal noise power in the base station. The total interference at the base station receiver for a user is the interference from the users in own cell and neighbouring cells (Outes 2001).

Solving the above equation for \(P_j\) gives,
The load factor $\rho_j$ of an $j^{th}$ active mobile is,

$$
P_j = \frac{I_{Total}}{1 + \left( \frac{E_b}{N_0} \right)_j v_j R_j w_c}
$$

The load factor $\rho_j$ of an $j^{th}$ active mobile is,

$$
\rho_j = \frac{I}{1 + \left( \frac{E_b}{N_0} \right)_j v_j R_j w_c}
$$

When a new user enters into the system, the call admission controller of the system first checks whether this connection requires real time service or non real time service. If it requires real time service, then it checks the resource availability in the home cell and also in the neighboring cells after reserving sufficient resources for the handoff calls. The reservation will be updated periodically and also at the time of new call arrival. If the system has enough resources, the call will be admitted otherwise the data rate of the non real time service calls will be reduced to the possible extent to accommodate the new call. If the new connection requires non real time service, this call will be admitted with full rate when the system has enough resource. If the system does not have enough resource, it will be admitted with half rate or even with quarter rate. The non real time call will be blocked only when the system can not provide the service even with this reduced rate. This will happen when the system is heavily loaded with more number of real time service calls. The resource requirement and availability are measured in terms of signal to interference ratio. Since call dropping is more annoying than call blocking
the handoff calls must be given higher priority. The reserved resources are used only by the handoff calls. Since the amount of resource reserved for handoff calls is based on the mobility estimation the resource utilization is also improved.

### 2.4.2 Fuzzy Model Based Call Admission Control

#### 2.4.2.1 Basic concepts of Fuzzy logic

In the Aristotelian logic a classical set can be defined as a set with a crisp boundary. For example, a classical set $A$ of real numbers greater than 6 can be expressed as

$$ A = \{x | x \geq 6 \} $$ (2.21)

There is a clear and unambiguous boundary “6”: if $x$ is greater than this number it belongs to the set $A$, otherwise $x$ does not belong to the set.

On the contrary, in the Fuzzy logic a set is defined without a crisp boundary. The transition from “belonging to the set” to “not belonging to the set” is gradual, thus representing the truth grade related to the definition of the concept. This smooth transition is characterized by the so-called ‘Membership Functions’ that give set flexibility in modelling commonly used linguistic expressions, like “the temperature is hot” or the “weather is warm”. A Fuzzy System consists of a Fuzzifier, an Inference Engine, a Fuzzy Rule Base and a Defuzzifier. The Fuzzifier transforms the values of the input parameters into the fuzzy linguistic terms through a set of Membership Functions. These fuzzy linguistic terms are the inputs of the Inference Engine, which will perform the logic inference according to the Fuzzy Rule Base. The Fuzzy Rule Base is constructed by the expert knowledge of the phenomenon.
(admission control, in this work). The Defuzzifier converts the results of the inference into the usable values for admission decisions. The Fuzzy Reasoning is also known as “approximate reasoning”, it is an inference procedure that derives conclusions from a set of fuzzy rules and known facts. It can be divided into four steps:

- Degrees of compatibility: compare the known facts with the antecedents of fuzzy rules to find the degrees of compatibility with respect to each antecedent Membership Function.

- Firing strength: combine degrees of compatibility with respect to antecedent Membership Functions in a rule using fuzzy AND or OR operators to form a firing strength that indicates the degree to which the antecedent part of the rule is satisfied.

- Qualified induced consequent Membership Functions: apply the firing strength to the consequent Membership Functions of a rule to generate a qualified consequent Membership Function.

- Overall output Membership Function: aggregate of all the qualified consequent Membership Functions to obtain an overall output Membership Function.

These four steps are employed in the fuzzy inference system shown in the following section.
2.4.2.2 Fuzzy system Model for CAC

Fuzzy Call Admission Control model used in our work is shown in the Figure 2.7.

![Fuzzy System Model Diagram]

**Figure 2.7 Fuzzy System Model**

The inputs parameters are,

\( S_T \) - Type of service request (Real time or Non-real time)

\( M \) - Mobility information of the new user.

\( I_{total} \) - Total interference in the cell, i.e. the inter-cell and intra-cell interference.

\( \rho \) - Load factor.
The corresponding fuzzy linguistic term set is,

\[ T(S_T) = \{ \text{high-RT, low-NRT} \} \]
\[ T(M) = \{ \text{low, medium, high} \} \]
\[ T(L) = \{ \text{low, medium, high} \} \]

The output linguistic variable, denoting the acceptability of the new call is,

\[ T(D): \{ \text{Strongly Rejected-SR, Weakly Rejected-WR, Weakly Accepted-WA, Strongly Accepted-SA} \}. \]

The relative membership functions are shown in Figure 2.8. The coefficients \( S_{Ta}, S_{Tb}, Ma, Mb, Mc, Md, La, Lb, Lc, Ld, Da, Db, Dc, Dd, De \) are the fuzzy set ranges of \( S_T, M, L \) and \( D \) respectively.

On the basis of the above fuzzy set, the Fuzzy Rule Base has been constructed and the rules for the admission criterion are listed in Table 2.1. Fuzzy inference algorithm is based on the Mamdani model.
Figure 2.8 Fuzzy membership functions
### Table 2.1  Fuzzy Rules

<table>
<thead>
<tr>
<th>Fuzzy Rule</th>
<th>Type of Service Request ($S_T$)</th>
<th>Mobility of the User (M)</th>
<th>Total Load Factor (L)</th>
<th>Fuzzy Decision (D)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>High</td>
<td>Low</td>
<td>Low</td>
<td>SA</td>
</tr>
<tr>
<td>2</td>
<td>High</td>
<td>Medium</td>
<td>Low</td>
<td>SA</td>
</tr>
<tr>
<td>3</td>
<td>High</td>
<td>High</td>
<td>Low</td>
<td>WA</td>
</tr>
<tr>
<td>4</td>
<td>High</td>
<td>Low</td>
<td>Medium</td>
<td>WA</td>
</tr>
<tr>
<td>5</td>
<td>High</td>
<td>Medium</td>
<td>Medium</td>
<td>WA</td>
</tr>
<tr>
<td>6</td>
<td>High</td>
<td>High</td>
<td>Medium</td>
<td>WA</td>
</tr>
<tr>
<td>7</td>
<td>High</td>
<td>Low</td>
<td>High</td>
<td>SR</td>
</tr>
<tr>
<td>8</td>
<td>High</td>
<td>Medium</td>
<td>High</td>
<td>SR</td>
</tr>
<tr>
<td>9</td>
<td>High</td>
<td>High</td>
<td>High</td>
<td>SR</td>
</tr>
<tr>
<td>10</td>
<td>Low</td>
<td>Low</td>
<td>Low</td>
<td>SA</td>
</tr>
<tr>
<td>11</td>
<td>Low</td>
<td>Medium</td>
<td>Low</td>
<td>SA</td>
</tr>
<tr>
<td>12</td>
<td>Low</td>
<td>High</td>
<td>Low</td>
<td>SA</td>
</tr>
<tr>
<td>13</td>
<td>Low</td>
<td>Low</td>
<td>Medium</td>
<td>SA</td>
</tr>
<tr>
<td>14</td>
<td>Low</td>
<td>Medium</td>
<td>Medium</td>
<td>WA</td>
</tr>
<tr>
<td>15</td>
<td>Low</td>
<td>High</td>
<td>Medium</td>
<td>WA</td>
</tr>
<tr>
<td>16</td>
<td>Low</td>
<td>Low</td>
<td>High</td>
<td>WA</td>
</tr>
<tr>
<td>17</td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td>WA</td>
</tr>
<tr>
<td>18</td>
<td>Low</td>
<td>High</td>
<td>High</td>
<td>WR</td>
</tr>
</tbody>
</table>
2.5 SUMMARY

The mobility based, parameters measurements based and fuzzy model based Call Admission Control and resource reservation schemes are described. In the mobility prediction approach, two different mobility models: one based on the LZW algorithm for data compression and the other based on the speed and direction of movement of the users, are used to predict not only the cells where the mobile users will handoff but also when it will handoff. In the parameters measurement based Call admission control scheme, the handoff dropping probability of the Class-2 calls is used as the parameter for adjusting the admission threshold to meet the QoS requirements of the multi class traffic. The class-3(best-effort service) traffic is serviced with the remaining resources, including those reserved but not being used by real-time service classes(both Class-1 and Class-2). In the wireless system, the decision made by the call admission controller is to be based on the imprecision and uncertain measurements due to user mobility, dynamic QoS requirements and varying channel conditions. Having the nature of coping with uncertainty and imprecision problems, fuzzy logic is expected to provide a good solution to the development of a call admission control scheme. In the fuzzy system model for CAC, the type of service request( Real time or non Real time), mobility information of the users and the current load factor of the system are used as the input parameters to make the admission decision.

In Chapter 3, simulation parameters used in the mobility based, parameters measurement based and fuzzy model based dynamic CAC schemes in WCDMA system are presented and the results obtained through the simulation are discussed.