Chapter 2

The Concept of Equivalent Interference in Multimedia CDMA Cellular Systems

In this chapter, equivalent interference estimators are proposed to estimate the capacity required for a call connection in WCDMA cellular systems. We first associate the equivalent interference with the traffic source characteristics and the radio resource defined by the power. Then, the equivalent interference estimator in the dedicated channels and shared channels are obtained, named EIE (equivalent interference estimator) and RRI (radio resource index), respectively. An analytical model is proposed and several techniques are adopted to obtain the EIE and the RRI. The EIE and the RRI can transform traffic parameters and quality-of-service (QoS) requirements of the call connection into a measure of resource in a unified metric, while keeping QoS requirements of existing calls guaranteed.

2.1 Notation List of Chapter 2

We summarize the important notations of this chapter in the following table. The temporary variables used in the proof which are also defined in the process of the derivation are not included in this table.

Table 2.1: Notation List of Chapter 2

Notation	Description
W	The bandwidth of the carrier of CDMA network
R_i	The transmission rate of connection i
SF_i	The spreading factor used by connection i
I_i	The resulting interference of connection i received at the base station in the uplink
I_v	The total received interference at base station in the uplink
$SIR_i(t)$	The received SIR connection i at frame time t
ρ_i	The ratio of the required SIR_i^* of connection i over the required SIR_v^* of voice connections
P_i	The required received power at base station for connection i
K	The number of traffic types in the considered system
N_k	The number of connections in traffic type k
$R_{p,i}$	The traffic parameter of peak rate for connection i
$R_{m,i}$	The traffic parameter of mean rate for connection i
$T_{p,i}$	The traffic parameter of peak rate duration for connection i
I_{th}	The interference threshold in the uplink direction pre-planned at base station
$P_{otg,k}^*$	The QoS requirement of outage probablity for traffic type k
$P_{D,i}^*$	The QoS requirement of packet dropping probablity for connection i
D_i^*	The QoS requirement of tolerable delay bound for connection i
SIR_i^*	The SIR requirement of connection i derived from the required E_b/N_0
$\tilde{r}_p$	The permission probability for connections using random access on the shared channels
$\tilde{r}_s$	The upper bounded random access successful probability

Notation	Description
Ω	The aggregated traffic process in the uplink
μ_Ω	The mean of aggregated traffic process in the uplink
σ_Ω	The variance of aggregated traffic process in the uplink
$\Lambda_X(\theta)$	The cumulant generating function of stochastic process $X(t)$ with parameter θ
$X_i^\tau(t)$	The point process of stochastic process $X(t)$ where the superscription τ indicates the arrival of traffic at time variable τ
$X_i^l(n)$	The length process of stochastic process $X(t)$ where the superscription l indicates the volume of the n-th arrival of traffic of which $n = X_i^\tau(t)$
$A_i(t)$	The arrival process of connection i
$B_i(t)$	The out process from connection i to the channel
$D_i(t)$	The departure process of connection i received at base station
$F_{n,i}(t)$	The failure process of connection i resulting due to excess delay or channel error, denoted by subscription $n = 1$ and $n = 2$, respectively
$H_p(t)$	The decision process of permission decision controller
$H_c(t)$	The decision process of channel

2.2 The Basic Concept of Equivalent Interference in Uplink Transmission

In this chapter, we propose the concept of equivalent interference in the uplink. We associate the equivalent interference with the traffic source characteristics and the radio resource defined by the power. In this section, we first show that the radio resource with specified physical layer requirements can be expressed in the concept of power. Then, the received interference is directly proportional to the required radio resource. Finally, the equivalent interference can be successfully related to the source characteristics and the requirements.

From many works on rate and power allocation, multi-code (variable rate) transmission, and resource allocation, we can conclude that the radio resource on the air can be defined by four factors: the coding scheme, the minimum spreading factor, the allowable power of the channel, and the number of the code channels. Based on this definition, then we may ask what is the basic unit to quantize the radio resource with these four associated parameters such that the physical layer requirement can be fulfilled. The following comes our proposed definition of radio resource unit.

Definition 1: Basic Resource Unit (BRU)

Assume that the maximum system load is planned by setting the maximum interference I_{th}. The BRU is defined as the total required received power to transmit one unit rate with loosest BER requirement on sufficient code channels using maximum spreading factor without channel coding under maximum system load condition. ■

This definition of BRU implies the smallest amount of resource to transmit measured by received power at base station side.

Property 2.1

In a given channel coding scheme, the required total received power to transmit a source in specific transmission rate with required BER is constant no matter how the combination of code channels and spreading factor as long as the required BRU is achieved.

Proof:

Denote R_i as the required source transmission rate of the considered connection, and SIR_i^* as the SIR requirement derived from the E_b/N_0 for adopted coding scheme. We decompose the source rate into arbitrary combination of N streams, namely $\{R_{i,1}, R_{i,2}, ..., R_{i,N}\}$, where N is a selected integer depending on the combination of the code channels. Let $P_{i,k}$ denote the required received power for stream k. Each stream k must satisfy the SIR_i^* requirement, that is,

$$\frac{P_{i,k}}{I} \cdot \frac{W}{R_{i,k}} \geq SIR_i^*. \tag{2.1}$$

The required received power for stream k is selected as

$$P_{i,k} = \frac{SIR_i^* \cdot I}{W} \cdot R_{i,k}. \tag{2.2}$$

Thus the total required received power for rate R_i is

$$\begin{aligned} P_i &= \sum_{k=1}^{N} P_{i,k} \\ &= \sum_{k=1}^{N} \frac{SIR_i^* \cdot I_v}{W} \cdot R_{i,k} \\ &= \frac{SIR_i^* \cdot I_v}{W} \cdot \sum_{k=1}^{N} R_{i,k} \\ &= \frac{SIR_i^* \cdot I_v}{W} \cdot R_i. \end{aligned} \tag{2.3}$$

We can find that the required power is only the function of required transmission rate and the SIR (or BER) requirement. Therefore, to transmit a source with specific rate is independent of the combination of code channels and the spreading factor. ■

When another channel coding scheme is used, such as convolutional code or turbo code, we can determine the total required received power of all used code channels to transmit one unit of rate by simply setting the corresponding $E_{b,i}/N_0$ for the coding scheme. In the same way, if the connection requires stricter BER requirement, we can also determine the required received power to transmit one unit of rate with different $E_{b,i}/N_0$. If there are N_{coding} possible coding schemes and N_{class} BER requirement classes, we can calculate the $BRU_{i,j} = \rho_{i,j} \cdot BRU_0$ for ith coding scheme and jth BER requirement in turn, where $\rho_{i,j}$ is the ratio between the required $SIR_{i,j}^*$ over the SIR_0^* which is the lowest SIR requirement of all traffic classes and coding schemes. For a connection using ith coding scheme and with jth BER requirement, we can determine the required amount of $BRU_{i,j}$ and finally the amount of BRU_0. Thus, different connections of different requirement will be measured in a unified and fair way.

In this manner, the radio resource parameters can be transform into the unit of power. The result in *Lemma 2* somehow explains several design in WCDMA cellular systems. In variable transmission rate case of *Power Management*, the allowable transmission power should be assigned for each connection to limit the load contributed by this connection and also match to the terminal radio access capability [49]. Moreover, the rate matching in physical layer in the uplink will select the combination of code channel and spreading factor on demand under the transmission power constraint. This BRU can be used as the unit for capacity credit at Node-B [50].

2.3 The Equivalent Interference Estimator for Uplink Transmission over Dedicated Channels

2.3.1 Introduction

The amount of radio resource required by a call connection in WCDMA is generally determined by its traffic parameters and QoS requirements. If a transformation that maps these parameters and requirements into a equivalent interference exists, it will be useful to radio resource management in WCDMA cellular systems [2].

The concept of equivalent interference for WCDMA cellular systems is similar to that of effective bandwidth for ATM networks. However, in the derivation of equivalent interference, the interference from both home cell and adjacent cells, the required power for each connection, and the random access behavior of MAC layer on the interference are to be considered. Moreover, the radio resource is constrained by a packet dropping ratio in order to guarantee the call-level requirement, while the bandwidth of an ATM node is a pre-defined hard capacity.

The algorithms to allocate radio resource has been studied in many literature [51], [52]. These algorithms just allocated power for each user to fulfill the required BER. From the QoS architecture in WCDMA [1], there are several requirements other than the BER requirement, such as packet error ratio and the delay, etc. If the allocated power considers only the BER, the call level requirement may not be satisfied to the time varying interference level. The characteristics of the interference process are influenced by the number and the characteristics of the active connections. In order to fulfill the BER and other call level requirements, a longer time scale interference process should be considered.

In this section, we derive the equivalent interference estimator, called by EIE, for the dedicated channels. In the dedicated channel, the radio resources are dedicated for the connections, and the transmission of each connection is in a transmission on demand fashion. Then, the interference process from each connection can be considered to be mutually independent and strictly stationary. Here, Gaussian approximation is adopted to derive the EIE in terms of the traffic parameters, the QoS requirements of SIR and outage probability. Real time and non-real time types of connections are considered.

2.3.2 System Model

In the considered model, K types of traffic are assumed. In the traffic type k, N_k connections are assumed. For differentiated bit-error-rate (BER) requirements set for the type-k traffic, we define their individual processing gains, denoted by SF_k. Note type-1 is the voice service which is selected to be the basic service class. All parameters will be expressed relative to the corresponding parameters for the type-1 connections. The SF_k is chosen to be the closest integer greater than the required spreading factor. Corresponding to each specific BER requirement and processing gain, the signal to interference ratio (SIR) threshold values of type-k traffic, denoted by SIR_k^* can be obtained. Several basic transmission rates (basic channels) are supported for connection i: (a) $R_i = R$, which is dedicated to active voice users (type-1) and is equal to the voice coding rate, and (b) $R_j = R \cdot \rho_j$, where connection j belongs to any traffic type-k, and the SIR requirement ratio $\rho_j = \frac{SIR_j^*}{SIR_i^*}$, and SIR_i^* is the SIR requirement for voice connections.

2.3.3 The Derivation of Equivalent Interference Estimator over Dedicated Channels

Denote the equivalent interference estimator (EIE) for type-k connections by $\hat{C}_k$, for $k = 1, \ldots K$. Let Ω be the aggregation traffic process of voice and data calls in the cell b, which can be expressed as

$$\Omega = \sum_{i=1}^{N_1} R_i + \sum_{k}^{K} \sum_{j=1}^{N_k} \cdot \rho_k \cdot R_j. \tag{2.4}$$

In order to fulfill the QoS requirements of outage probabilities, the process Ω should satisfy the following constraints

$$Pr\{\Omega > 1/SIR_k^*\} < \min\left\{P_{otg1}^*, \ldots, P_{otgK}^*\right\} \tag{2.5}$$

with constraints on the number of users as

$$N_1\hat{C}_1 + \sum_{k=2}^{K} N_k\hat{C}_k < \frac{1}{SIR_1^*}, \tag{2.6}$$

$$N_1\hat{C}_1 + \sum_{k=2}^{K} N_k\hat{C}_k < \frac{\rho_k}{SIR_k^*}. \tag{2.7}$$

The EIE can be derived by solving the above inequalities. In this subsection, we take 2 types, $K = 2$, for example. Assume that Ω possesses Gaussian property and its mean and the variance consists of the mean and the variance of each individual connection, respectively. $R_{m,1}$ ($R_{m,2}$) and σ_1 (σ_2) are the mean and the variance of the rate generated by a type-1 (type-2) call. Since the type-1 and type-2 source models are herein assumed to be an on-off process and a batch Poisson process, respectively, the σ_1 and σ_2 can be obtained by

$$\sigma_1^2 = (R_{p,1} - R_{m,1})R_{p,1}, \tag{2.8}$$

$$\sigma_2^2 = \frac{R_G^2 \cdot R_{m,2}^2}{(T_{p,2} - 1)}, \tag{2.9}$$

where $R_{p,1}$ is peak rate of a voice call and $T_{p,2}$ is the peak rate duration of a data call. The mean μ_Ω and the variance σ_Ω of Ω are then

$$\mu_\Omega = N_1 \cdot R_{m,1} + N_2 \cdot R_{m,2}, \tag{2.10}$$

$$\sigma_\Omega^2 = N_1 \cdot \sigma_1^2 + N_2 \cdot \sigma_2^2. \tag{2.11}$$

By normalizing the variable Ω, (2.5) can be written as

$$P\{\Omega' > \frac{(1/SIR_1^*) - \mu_\Omega}{\sigma_\Omega}\} < P_{otg1}^*, \tag{2.12}$$

where Ω' is a normalized Gaussian random variable. Let β_1 be a constant such that $Q(\beta_1) = P_{otg1}^*$. As (2.12) will be hold, the condition

$$\frac{(1/SIR_1^*) - \mu_\Omega}{\sigma_\Omega} > \beta_1 \tag{2.13}$$

must be satisfied. By substituting (2.10) and (2.11) into (2.13), the constraint for type-1 traffic becomes

$$(1/SIR_1^*) - N_1 \cdot R_{m,1} + N_2 \cdot R_{m,2} > \beta_1 \cdot (N_1 \cdot \sigma_1^2 + N_2 \cdot \sigma_2^2)^{1/2}. \tag{2.14}$$

Similarly, the constraint for type-2 traffic can be obtained by

$$(R_G/SIR_2^*) - N_1 \cdot R_{m,1} + N_2 \cdot R_{m,2} > \beta_2 \cdot (N_1 \cdot \sigma_1^2 + N_2 \cdot \sigma_2^2)^{1/2}, \tag{2.15}$$

where β_2 is a constant such that $Q(\beta_2) = P_{otg2}^*$. Consequently, $\hat{C}_1$ and $\hat{C}_2$ can be obtained from the four constraints as given in (2.14), (2.15), (2.6), and (2.7).

To derive the $\hat{C}_i$ in the more traffic type case, bandwidth allocation is required to define the resource partition policy among the different types of connections, and linear or nonlinear programming techniques can be used.

2.4 The Equivalent Interference Estimator for Uplink Transmission over Shared Channels

2.4.1 Introduction

In this section, we study the equivalent interference estimator for connections using shared channel with delay bound in WCDMA cellular systems. Since the equivalent interference indicates the radio resource at base station side used by a connection, the equivalent interference estimator in shared channel is named *radio resource index* hereafter.

We define BER as the packet-level requirement and the packet dropping ratio and the tolerable delay as the call level requirements. The radio resource allocation algorithm considers both the packet-level requirement and the call level requirements. The RRI is derived for uplink call connection as an indication for base station to estimate the amount of radio resource used by the user at receiver side. We first define analytical elements in a model to equivalently describe the behavior of the system. Some properties are developed to derive the input and output relationship of the analytical elements. We then form the failure process consisted of all the dropped packets due to either excess delay in the transmitter or the channel error. In order to fulfill the BER and dropping ratio requirement, we can find that the required radio resource, converted into the unit of power, is in terms of the SIR, source traffic characteristics and call level requirements. Based on the concept of multiplexing different users on the shared channel, we finally calculate the equivalent radio resource index consumed by each user at the base station side. The RRI performs as a function mapping from a parametric space, which is constituted by traffic parameters and QoS requirements, to the metric space, which is of unified metric to each different connections, while keep QoS requirements of all existing calls guaranteed.

2.4.2 System Model

We assume the connection with bursts of which packets of a burst arrive in a batch fashion. The head-of-line packet of a burst is firstly sent as a request for resource reservation, where the request permission probability $\tilde{r}_p$ is determined and broadcasted by the radio network controller. If the first packet is not permitted to transmit in this frame or its transmission is corrupted in the air interface, it will retry. If the first packet fails to transmit successfully or be acknowledged before the maximum tolerable delay, the whole burst will be dropped. Once the first packet is successfully acknowledged, the remaining packets of burst can be sequentially sent without any further request. The packet sent out but corrupted in the air interface would be discarded by the receiver. We set a SIR threshold, SIR_0, to determine if the packet is successfully received or not. We can set the maximum system load I_{th}, and the required received power to transmit information in a basic rate with required SIR is P_i^0. And we assume that the call-level QoS requirements of user i are the packet dropping ratio, $R^*_{D,i}$, and the maximum tolerable delay, D^*_i.

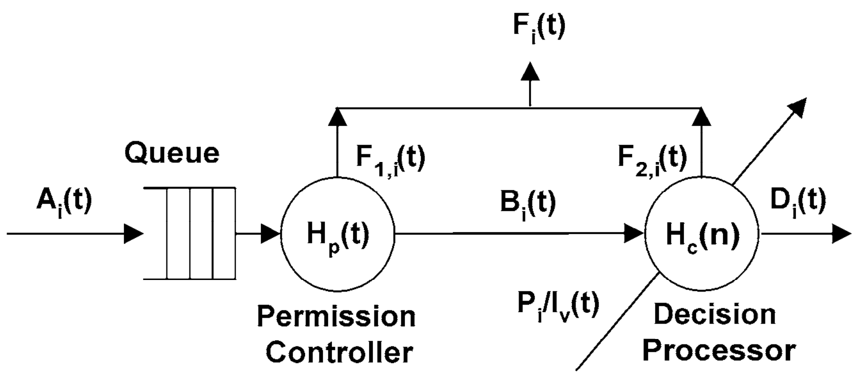

Figure 2.1: The analytical model of the uplink connection i in WCDMA.

There are N_v connections in the uplink of WCDMA cellular systems. The analytical model for uplink connection i is shown in Fig. 2.1, where the capital letter X is to represent the cumulative process of a process which is denoted by small letter x in the following. The $A_i(t)$ denotes the *arrival process* of user i. The $H_p(n)$ is the *access process* for arrivals in permission controller, where packets are transmission permitted or dropped. The dropped packets are directed toward a *failure process* denoted by $F_{1,i}(t)$; the permitted packets form an *output process* denoted by $B_i(t)$. Subsequently, the process $B_i(t)$ enters into the *decision process* denoted by $H_c(n)$. The decision processor determines the way the packet goes in the light of the level of *aggregated interference process* denoted by $I_v(t)$, which is the summation of the adjacent-cell interference $I_a(t)$ and the home-cell interference $I_h(t)$. The transmission power P_i of the transmission packet of connection i forms a process $I_i(t)$ on the air interface, which indicates the power component contributed by connection i onto the transmission channel. If $P_i/I_v(t)$ is less than SIR_0, the packet will be directed to a *failure process* $F_{2,i}(t)$, otherwise, the packet will be successfully received and forms a *departure process*, denoted by $D_i(t)$. Note that for convenience, $I_h(t) = \sum_{j=1}^{N_v} I_j(t)$ including that of user i is assumed, which would be the upper-bounded interference for each user.

2.4.3 The Design of Radio Resource Index for Shared Channels

The arrival process $A_i(t)$ can be decomposed into a point process $A_i^{\tau}(t)$ and a burst length process $A_i^l(n)$. The compound process has the property described in lemma 1.

Theorem 1:

For a given *cumulant generating function* $\Lambda_{I_v}(\theta)$, and θ^* to be the solution to $\Lambda_{I_v}(\theta)/\theta = I_{th}$, the received power for the basic rate transmission of connection i, required to fulfill $R_{D,i}^*$ and SIR_i, should be set to

$$P_i \geq \frac{-log(1 - \frac{1-R_{D,i}^*}{\tilde{r}_s}) + \Lambda_{I_v}(\theta^*)}{\theta^*} \cdot SIR_i, \qquad (2.16)$$

where $\tilde{r}_s$ denotes the successful access probability in the permission controller. The RRI $\hat{C}_i$ of connection i, in terms of the cumulant generating function of input process $\Lambda_{A_i}(\theta)$ and the received power P_i, is given by

$$\hat{C}_i = \frac{\Lambda_{I_i}(\theta^*)}{\theta^*} = \frac{\Lambda_{A_i}(P_i \cdot \theta^*) - \Lambda_{A_i}(P_i \cdot \theta^*(1-\tilde{r}_s))}{\theta^*}. \tag{2.17}$$

Also, the packet dropping ratio of each connection can be guaranteed if the following constraint is satisfied, $\sum_i \hat{C}_i \leq \frac{\Lambda_{I_v}(\theta^*)}{\theta^*}$. ■

The ratio $\frac{\Lambda_{I_i}(\theta)}{\theta}$ varies from the mean rate to the peak rate of the process as θ goes from 0 to ∞. By the *Taylor* expansion of the cumulant generating function, θ^* is the best weighting of each existing order of the moments of the process and depends on the setting of the system. Several lemmas are presented before the above results can be derived.

Lemma 2:

The arrival process $A_i(t)$, consisting of a point process $A_i^{\tau}(t)$ and a length process of the n-th burst $A_i^{l}(n)$, has the cumulant generating function, denoted by $\Lambda_{A_i}(\theta)$, which can be expressed as

$$\Lambda_{A_i}(\theta) = \Lambda_{A_i^{\tau}}(\Lambda_{A_i^{l}}(\theta)). \tag{2.18}$$

The derivation is given in the Appendix A. ■

Lemma 3:

Consider the point process $A_i^{\tau}(t)$ through the permission controller with access process$H_p(n)$. Let $\Lambda_{H_p}(\theta)$ be the cumulant generating function of the $H_p(n)$. The point process output from the permission controller for $A_i^{\tau}(t)$, denoted by $B_i^{\tau}(t)$, has the cumulant generating function $\Lambda_{B_i^{\tau}}(\theta)$ obtained by

$$\Lambda_{B_i^{\tau}}(\theta) = \Lambda_{A_i^{\tau}}(\Lambda_{H_p}(\theta)). \tag{2.19}$$

The derivation is given in the Appendix A. ■

The burst length process $B_i^l(n)$ of the output process $B_i(t)$ is preserved the same as $A_i^l(n)$. The cumulant generating function of $B_i(t)$, denoted by $\Lambda_{B_i}(\theta)$, can be directly written as

$$\begin{aligned}\Lambda_{B_i}(\theta) &= \Lambda_{B_i^\tau}(\Lambda_{B_i^l}(\theta)) \\ &= \Lambda_{B_i^\tau}(\Lambda_{A_i^l}(\theta)).\end{aligned} \tag{2.20}$$

Lemma 4:

The arrival process $A_i(t)$ is divided into two sub-flows: $B_i(t)$ and $F_{1,i}(t)$. And the cumulant generating function of $F_{1,i}(t)$ can be expressed as

$$\Lambda_{F_{1,i}}(\theta) = \Lambda_{A_i}(\theta) - \Lambda_{B_i}(\theta). \tag{2.21}$$

The derivation is given in Appendix A. ■

Lemma 5:

The outage probability of connection i, denoted by R_{otg}, related to the aggregated interference can be obtained by

$$R_{otg} = \lim_{N_v \to \infty} \mathbf{Pr}[I_v \in G_i] \approx e^{-\Lambda^*_{I_v}(G_i)}, \tag{2.22}$$

where $G_i = \{I_v | P_i/I_v < SIR_0\}$, $\Lambda^*_{I_v}(G_i)$ is the rate function of *interference process*, and N_v is the total number of calls in WCDMA. The derivation is obtained in the Appendix A. ■

Lemma 6:

The cumulant generating function of the decision process $H_c(n)$, denoted by $\Lambda_{H_c}(\theta)$, can be derived as

$$\Lambda_{H_c}(\theta) = log(e^{\theta - \Lambda^*_{I_v}(G_i)} + 1 - e^{-\Lambda^*_{I_v}(G_i)}). \tag{2.23}$$

The derivation is given in the Appendix A. ■

The failure process $\Lambda_{F_{2,i}}(\theta)$ from the decision processor can be gotten by

$$\Lambda_{F_{2,i}}(\theta) = \Lambda_{B_i}(\Lambda_{H_c}(\theta)). \tag{2.24}$$

Similar to Eq. (2.21), $\Lambda_{D_i}(\theta)$ can be yielded as

$$\Lambda_{D_i}(\theta) = \Lambda_{B_i}(\theta) - \Lambda_{F_{2,i}}(\theta). \tag{2.25}$$

Based on these results derived before, we can predict the packet dropping ratio for each connection. For a given packet dropping ratio requirement $R^*_{D,i}$, we can obtain a channel condition constraint by the following lemma.

Lemma 7:

For the QoS requirement $R^*_{D,i}$, the aggregated interference process $I_v(t)$ on connection i has the constraint given by

$$\Lambda^*_{I_v}(G_i) \geq -log(1 - \frac{1 - R^*_{D,i}}{\tilde{r}_s}), \tag{2.26}$$

where $\tilde{r}_s$ is the successful probability of a burst from the permission controller. ■

Proof of Theorem 1:

Proof: The ratio $\frac{\Lambda_{I_i}(\theta)}{\theta}$ is the equivalent power of connection i received at base station. This equivalent power generated by each connection is regarded as the equivalent load on the system. From the superposition property of $\frac{\Lambda_{I_v}(\theta)}{\theta} = \sum_i \frac{\Lambda_{I_i}(\theta)}{\theta}$, the total equivalent power $\frac{\Lambda_{I_v}(\theta)}{\theta}$ should be constrained by the maximum system load I_{th}, to guarantee the packet dropping ratios of all existing connections. Define $\mathcal{G}_\theta = \{\theta : \Lambda_{I_v}(\theta)/\theta \leq I_{th}\}$. Since $\Lambda_{I_v}(\theta)$ is convex and $\Lambda_{I_v}(\theta)/\theta$ is increasing, $\Lambda_{I_v}(\theta)/\theta \leq \Lambda'_{I_v}(\theta)$ for $\theta \in \mathcal{G}_\theta$ and θ^* is the point in $\mathcal{G}_\theta$ that maximizes $\left\{\frac{P_i^0}{SIR_i} \cdot \theta - \Lambda_{I_v}(\theta)\right\}$ with $\frac{P_i^0}{SIR_i} \leq I_{th}$. Also, owing to the equal received power control, $\frac{P_i^0}{SIR_i}$ and θ^* are the same for all i. Additionally,

$$\Lambda_{I_v}^*(G_i) = \inf_{\alpha \in G_i} \left(\Lambda_{I_v}^*(\alpha)\right) = \sup_{\theta \in \mathcal{G}_\theta} \left\{ \frac{P_i}{SIR_i} \cdot \theta - \Lambda_{I_v}(\theta) \right\}, \tag{2.27}$$

where $\frac{P_i}{SIR_i} \geq \frac{P_i^0}{SIR_i}$. Let θ_i^* be the solution to $\frac{P_i}{SIR_i} = \Lambda'_{I_v}(\theta)$. Since $\Lambda'_{I_v}(\theta)$ is increasing, $\theta_i^* \geq \theta^*$ for all i. Therefore, θ^* is the maximum point in $\mathcal{G}_\theta$ with the largest $\left(\frac{P_i}{SIR_i} \cdot \theta - \Lambda_{I_v}(\theta)\right)$, and (2.27) becomes $\Lambda_{I_v}^*(G_i) = \frac{P_i}{SIR_i} \cdot \theta^* - \Lambda_{I_v}(\theta^*)$. From (2.A.15), $\frac{P_i}{SIR_i} \cdot \theta^* - \Lambda_{I_v}(\theta^*) \geq -log(1 - \frac{1-R_{D,i}^*}{\tilde{r}_s})$, and then (2.16) can be obtained.

The $\hat{C}_i$ is defined for connection i as the equivalent power generated by connection i, $\frac{\Lambda_{I_i}(\theta^*)}{\theta^*}$. Since $I_i(t) = P_i \cdot b_i(t)$, and using *Lemma 1* to *Lemma 3*, the $\hat{C}_i$ is given by $\hat{C}_i = \frac{\Lambda_{I_i}(\theta^*)}{\theta^*} = \frac{\Lambda_{B_i}(P_i\theta^*)}{\theta^*}$, and (2.17) is obtained. If $\sum_i \hat{C}_i$ is less than the $\Lambda_{I_v}(\theta^*)/\theta^*$, then the QoS requirements of each connection can be guaranteed. ■

The RRI $\hat{C}_i$ of connection i indicates the effective load contributed by the user in the multiplexing transmission environment. As the call requirement becomes stricter, the required power is scaled up, the the resource consumed by this user is therefore increased. The θ^* is the scaling factor such that $\hat{C}_i$ varies from the mean rate (as θ^* approaches 0) to the peak rate (as θ^* approaches ∞). From the Taylor's expansion of $\hat{C}_i$, θ^* is the weighting of higher order statistics. As the value of θ^* is selected smaller, the estimated resource for each connection is less, but the resource marginal C^*/θ^* for keeping the dropping probability fulfilled is larger. Thus, the total available resource inside the cell depends on the adjacent cell interference process, the scaling factor θ^*, and the channel condition constraint.

2.4.4 Results and Conclusions

In this section, the *Radio Resource Index* is obtained to estimate the required radio resource for a call connection. We investigate the characteristics of uplink transmission in WCDMA cellular systems, which make the derivation of RRI different from the effective bandwidth in ATM networks. Based on the properties of the up-link transmission in WCDMA, we model the behavior of the uplink transmission, and derive the input/output process relationships for each element in the model by using large deviation techniques. We finally derive the expression for the RRI for each connection and examine the physical meanings of RRI.

To verify the effectiveness of the proposed RRI, we examine if the packet dropping requirement of each connection is satisfied under the way the resource is allocated in light and heavy load conditions. The heavier the load is (i.e. the summation of RRI of each connection is closer to the maximum value), the closer the measured packet dropping ratio to the requirement is. In the worst case, that is, the summation of RRI of each connection is equal to the maximum value, we expect that the measured packet dropping ratio of each connection will approximate it's requirement. Two scenarios are presented here. In the first scenario, the only traffic is 12.2k conversational service with $R^*_{D,i} = 0.02$. In the second scenario, there are three types of traffic: 12.2k conversational service with $R^*_{D,i} = 0.02$, 8.6k conversational service with $R^*_{D,i} = 0.02$, and 4.75k conversational service with $R^*_{D,i} = 0.05$. In the simulations, we set SIR_0 as $-14dB$, I_{th}/P_i^0 as 52, the permission probability $\tilde{r}_p$ as 0.9, and θ^* as 1.1.

Table 2.2 shows the RRI of each connection, the simulated and theoretical packet dropping ratio, and the theoretical maximum number of users in one cell in both scenario 1 (a) and scenario 2 (b). In the scenario 1, the simulated packet dropping ratio $\hat{R}_D$ is 2.5 times smaller than the theoretical result $\hat{R}^*_D$. As the number of users approaches 54, the $\hat{R}_D$ is about 0.02. This indicates the proposed RRI is conservative in resource allocation to protect connections against the interference, and attains about 88 percentage of best achievable resource utilization efficiency. In the scenario 2, we set different packet dropping ratio requirements and different transmission rate for different type of connections,

Table 2.2: The theoretical and simulation results.

Scenario 1				
Parameters	$\hat{C}_i$	$\hat{R}_D$	$\hat{R}_D^*$	number of users
Single type	1.08	0.00792	0.02	48

(a) Scenario 1 with single traffic types

Scenario 2				
Parameters	$\hat{C}_i$	$\hat{R}_D$	$\hat{R}_D^*$	number of users
Type 1	1.08	0.00626	0.02	25
Type 2	0.67	0.00621	0.02	25
Type 3	0.31	0.0109	0.05	27

(b) Scenario 2 with three traffic types

and the RRI for each type of connection is shown in Table 2.2 (b). The simulated packet dropping ratios of type-1 and type-2 connections are about 3 times less than the theoretical values, and the simulated value of type-3 connections is almost 5 times smaller than the theoretical values. The packet dropping ratio of type 3 connections is higher than those of type-1 and type-2 connections because looser packet dropping ratio requirement is set for type 3 connections and the allocated power can be decreased. The proposed RRI can allocate proper radio resource to the connection according to not only the traffic characteristics but also the specific requirement. The RRI is a flexible and simple mapping from the traffic parameters and the QoS requirements in the situation of any number of types of connections and requirements.

Appendix: Proof of Lemmas

Proof of Lemma 1:

Express $A_i(t)$ as

$$A_i(t) = \sum_{n=1}^{\infty} 1_{A_i^{\tau}(t)}(n) \cdot a_i^l(n), \tag{2.A.1}$$

where

$$1_{A_i^\tau(t)}(n) = \begin{cases} 1, & n \le A_i^\tau(t) \\ 0, & n > A_i^\tau(t) \end{cases}. \tag{2.A.2}$$

And assume the arrival process stationary. Then we have

$$\begin{aligned}
\Lambda_{A_i}(\theta) &= \lim_{t\to\infty} \frac{1}{t} log E\left[e^{\theta A_i(t)}\right] \\
&= \lim_{t\to\infty} \frac{1}{t} log E\left[e^{\theta \sum_{n=1}^{\infty} 1_{A_i^\tau(t)}(n)\cdot a_i^l(n)}\right] \\
&= \lim_{t\to\infty} \frac{1}{t} log \sum_{k=1}^{\infty} E\left[e^{\theta \sum_{n=1}^{A_i^\tau(t)} 1_{A_i^\tau(t)}(n)\cdot a_i^l(n)} | A_i^\tau(t) = k\right] \cdot P\left[A_i^\tau(t) = k\right] \\
&= \lim_{t\to\infty} \frac{1}{t} log \sum_{k=1}^{\infty} e^{\Lambda_{A_i^l}(\theta)\cdot k} \cdot P\left[A_i^\tau(t) = k\right] \\
&= \lim_{t\to\infty} \frac{1}{t} log E\left[e^{\Lambda_{A_i^l}(\theta)\cdot A_i^\tau(t)}\right] \\
&= \Lambda_{A_i^\tau}(\Lambda_{A_i^l}(\theta)).
\end{aligned} \tag{2.A.3}$$

■

Proof of Lemma 2:

Define $\tilde{r}_s$ as the successful access probability of a burst, and obtain $\tilde{r}_s$ can be obtained by

$$\tilde{r}_s = 1 - (1 - \tilde{r}_p(1 - R_{otg}))^{D_i^*} \tag{2.A.4}$$

For the permission controller, it can be expressed as

$$h_p(a_i^\tau(t)) = \begin{cases} 1_p(u(n)), & a_i^\tau(t) = 1 \\ 0, & otherwise \end{cases}, \tag{2.A.5}$$

where

$$1_p(u(n)) = \begin{cases} 1, & u(n) \le \tilde{r}_s \\ 0, & u(n) > \tilde{r}_s \end{cases},$$

and $u(n)$ is a random variable uniformly distributed in the region $[0, 1]$. We can obtain the $\Lambda_{H_p}(\theta)$ as

$$\Lambda_{H_p}(\theta) = log\left(\tilde{r}_s \cdot e^{\theta} + 1 - \tilde{r}_s\right). \tag{2.A.6}$$

The cumulant generating function of $B_i^\tau(t)$ can be derived as

$$
\begin{aligned}
\Lambda_{B_i^\tau}(\theta) &= \lim_{t\to\infty} \frac{1}{t} logE\left[exp\{\theta(H_p(A_i^\tau(t)))\}\right] \\
&= \lim_{t\to\infty} \frac{1}{t} logE\left[E\left[exp\{\theta(H_p(k))\}|A_i^\tau(t)=k\right]\right] \\
&= \lim_{t\to\infty} \frac{1}{t} logE\left[exp\{k \cdot \lim_{k\to\infty} \frac{1}{b} logE\left[exp\{\theta(H_p(k))\}|A_i^\tau(t)=k\right]\}\right] \\
&= \lim_{t\to\infty} \frac{1}{t} logE\left[exp\{\Lambda_{H_p}(\theta) \cdot A_i^\tau(t)\}\right] \\
&= \Lambda_{A_i^\tau}(\Lambda_{H_p}(\theta)) \\
&= \Lambda_{A_i^\tau}\left(log\left(\tilde{r}_s \cdot e^\theta + 1 - \tilde{r}_s\right)\right). \qquad (2.A.7)
\end{aligned}
$$

■

Proof of Lemma 3:

The arrival process $A_i(t)$ is divided into of $B_i(t)$ and $F_{1,i}(t)$, that is $A_i(t) = B_i(t) + F_{1,i}(t)$. Thus, the cumulant generating function of $A_i(t)$ is

$$
\begin{aligned}
\Lambda_{A_i}(\theta) &= \lim_{t\to\infty} \frac{1}{t} logE\left[e^{\theta A_i(t)}\right] \\
&= \lim_{t\to\infty} \frac{1}{t} logE\left[e^{\theta(B_i(t)+F_{1,i}(t))}\right] \\
&= \Lambda_{B_i}(\theta) + \Lambda_{F_{1,i}}(\theta). \qquad (2.A.8)
\end{aligned}
$$

Therefore,

$$\Lambda_{F_{1,i}}(\theta) = \Lambda_{A_i}(\theta) - \Lambda_{B_i}(\theta).$$

■

Proof of Lemma 4:

$\Lambda^*_{I_v}(G_i)$ is the rate function on the event G_i for process $I_v(t)$, and is expressed by the *Legendre Transform* [5] of its cumulant generating function, $\Lambda^*_{I_v}(G_i) = \inf_{\alpha\in G_i}\{\sup_\theta[\theta\cdot\alpha - \Lambda_{I_v}(\theta)]\}$. The result of the unbuffered resources model in [55] states that if the aggregation of random variables at time t, denoted by $S(t)$, exceeds the capacity Z, then the large deviation approximation to the probability measure on the resource overflow event "$S(t) \geq Z$" is

$$\mathbf{Pr}[S(t) \geq Z] \approx exp\left\{\inf_{\theta}\left\{\sum_{j=1}\Lambda_j(\theta) - \theta Z\right\}\right\}, \tag{2.A.9}$$

where $\Lambda_j(\theta)$ is the cumulant generating function of source j. The resource overflow event is regarded as the event "$I_v(t) \in G_i$" in the analysis, and the right-hand side of (2.A.9) is the rate function that measures the set G_i. Therefore, the outage probability R_{otg} can be obtained by (2.22). ■

Proof of Lemma 5:

The decision process of decision processor can be modelled as $r_c(t) = 1_{G_i}(I_v(t))$, where

$$1_{G_i}(I_v(t)) = \begin{cases} 1, & I_v(t) \in G_i, \\ 0, & otherwise. \end{cases}$$

Then $\Lambda_{H_c}(\theta)$ can be obtained by

$$\begin{aligned} \Lambda_{H_c}(\theta) &= \lim_{n\to\infty} \frac{1}{n} logE\left[e^{\theta H_c(n)}\right] \\ &= \lim_{n\to\infty} \frac{1}{n} logE\left[e^{\theta \sum_{k=1}^{n} 1_{G_i}(I_v(\tau(k)))}\right] \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} logE\left[e^{\theta 1_{G_i}(I_v(\tau(k)))}\right] \\ &= \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} log(R_{otg}e^{\theta} + (1 - R_{otg})) \\ &= log(R_{otg}e^{\theta} + (1 - R_{otg})) \\ &= log(e^{\theta - \Lambda^*_{I_v}(G_i)} + 1 - e^{-\Lambda^*_{I_v}(G_i)}). \end{aligned} \tag{2.A.10}$$

■

Proof of Lemma 7: $< pf >$

The cumulant generating function for a general process $Z(t)$, denoted by $\Lambda_Z(\theta)$, has two properties. As $\theta \to 0$, it equals to zero and its derivative equals to the mean of the process μ_z.

The packet dropping ratio of connection i can be derived as

$$\hat{R}_D(i) = \lim_{t\to\infty} \frac{1}{\mu_{A_i}} \Lambda'_{F_{1,i}(t)}(\theta)\Big|_{\theta=0} + \lim_{t\to\infty} \frac{1}{\mu_{A_i}} \Lambda'_{F_{2,i}(t)}(\theta)\Big|_{\theta=0}. \quad (2.A.11)$$

For $\Lambda'_{F_{1,i}(t)}(\theta)$ in the first term of (2.A.11), it can be obtained by

$$\begin{aligned}
&\Lambda'_{F_{1,i}(t)}(\theta)\Big|_{\theta=0}\\
=\ & \frac{d\Lambda_{A_i}(\theta)}{d\theta}\Big|_{\theta=0} - \frac{d\Lambda_{A_{\tau,i}}(\Lambda_{H_p}(\Lambda_{A_i^l}(\theta)))}{d\theta}\Big|_{\theta=0}\\
=\ & \frac{d\Lambda_{A_i}(\theta)}{d\theta}\Big|_{\theta=0} - \frac{d\Lambda_{A_{\tau,i}}(\theta_1)}{d\theta_1}\Big|_{\theta_1=\Lambda_{H_p}(\Lambda_{A_i^l}(\theta))} \cdot \frac{d\Lambda_{H_p}(\theta_2)}{d\theta_2}\Big|_{\theta_2=\Lambda_{A_i^l}(\theta)} \cdot \frac{d\Lambda_{A_i^l}(\theta)}{d\theta}\Big|_{\theta=0}\\
=\ & \frac{d\Lambda_{A_i}(\theta)}{d\theta}\Big|_{\theta=0} - \frac{d\Lambda_{A_{\tau,i}}(\theta_1)}{d\theta_1}\Big|_{\theta_1=0} \cdot \frac{d\Lambda_{H_p}(\theta_2)}{d\theta_2}\Big|_{\theta_2=0} \cdot \frac{d\Lambda_{A_i^l}(\theta)}{d\theta}\Big|_{\theta=0}\\
=\ & \mu_{A_i} - \mu_{A_{\tau,i}} \cdot \tilde{r}_s \cdot \mu_{A_i^l}, \qquad (2.A.12)
\end{aligned}$$

and $\tilde{r}_s$ is given in Eq. (2.A.4). For $\Lambda'_{F_{2,i}(t)}(\theta)$ in the second term of (2.A.11), it can be derived as

$$\begin{aligned}
&\Lambda'_{F_{2,i}(t)}(\theta)\Big|_{\theta=0}\\
=\ & \frac{d\Lambda_{B_i}(\Lambda_{H_c}(\theta))}{d\theta}\Big|_{\theta=0}\\
=\ & \frac{d\Lambda_{B_i}(\theta_1)}{d\theta_1}\Big|_{\theta_1=0} \cdot \frac{d\Lambda_{H_c}(\theta)}{d\theta}\Big|_{\theta=0}\\
=\ & \frac{d\Lambda_{B_i}(\theta_1)}{d\theta_1}\Big|_{\theta_1=0} \cdot \lim_{n\to\infty} \frac{1}{n} \cdot \frac{E\left[H_c(t)e^{\theta H_c(n)}\right]}{E\left[e^{\theta H_c(n)}\right]}\Big|_{\theta=0}\\
=\ & \frac{d\Lambda_{B_i}(\theta_1)}{d\theta_1}\Big|_{\theta_1=0} \cdot \lim_{n\to\infty} E\left[\frac{H_c(n)}{n}\right]\\
=\ & \frac{d\Lambda_{B_i}(\theta_1)}{d\theta_1}\Big|_{\theta_1=0} \cdot \lim_{n\to\infty} E\left[\frac{\sum_{k=1}^{n} 1_{G_i}(I_v(\tau(n)))}{n}\right]\\
=\ & \frac{d\Lambda_{B_i}(\theta_1)}{d\theta_1}\Big|_{\theta_1=0} \cdot E\left[I_v(\tau(n))\right]\\
=\ & \mu_{B_i} \cdot exp\{-\Lambda^*_{I_v}(G_i)\}. \qquad (2.A.13)
\end{aligned}$$

Therefore, $\hat{R}_D(i)$ can be rewritten as

$$\hat{R}_D(i) \quad = \quad \mu_{A_i} - \mu_{A_{\tau,i}} \cdot \tilde{r}_s \cdot \mu_{A_i^l} + \mu_{B_i} \cdot exp\{-\Lambda^*_{I_v}(G_i)\}. \qquad (2.A.14)$$

Consequently, given the QoS constraint $\hat{R}_D(i) \leq R^*_{D,i}$, $\Lambda^*_{I_v}(G_i)$ can be reformatted and obtained by

$$\Lambda^*_{I_v}(G_i) \geq -log(1 - \frac{1 - R^*_{D,i}}{\tilde{r}_s}). \qquad (2.A.15)$$

This result shows us that the function $\Lambda^*_{I_v}(G_i)$ measuring the set G_i should satisfy condition such that the outage probability faced by this user can fulfill the packet dropping ratio. As $1 - R^*_{D,i}$ approaches $\tilde{r}_s$, that is the packet dropping due to channel loss, the required power $\Lambda^*_{I_v}(G_i)$ should be increased. ■