Wireless Netw (2011) 17:231–246
DOI 10.1007/s11276-010-0275-x

Optimal rate allocation and QoS-sensitive admission control in wireless integrated networks

Hyun M. Shin · Chae Y. Lee

Published online: 31 August 2010
© Springer Science+Business Media, LLC 2010

Abstract The problem of Call Admission Control and rate allocation in loosely coupled wireless integrated networks is investigated. The related Radio Resource Management schemes were introduced to improve network performance in wireless integrated networks. However, these schemes did not reflect the independence and competitiveness of loosely coupled wireless integrated networks. Furthermore, given that users have different requirements for price and Quality of Service (QoS), they are able to select a network according to their preference. We consider a scenario with two competitive wireless networks, namely Universal Mobile Telecommunications System cellular networks and Wireless Local Area Networks. Users generate two types of traffic with different QoS requirements: real-time and non-real-time. We propose a scheme that exploits a mathematical model for the control of call admission and adopt a noncooperative game theory-based approach to address the rate allocation problem. The purpose is to maximize the revenue of the network providers while guaranteeing a level of QoS according to user needs. Simulation results show that the proposed scheme provides better network performance with respect to packet loss rate, packet delay time, and call-blocking probability than other schemes when the data rates are allocated to each call at the point that maximizes the revenue of network providers. We further demonstrate that a Nash equilibrium always exists for the considered games.

Keywords Wireless integrated networks · Call admission control · Rate allocation · Noncooperative game theory

H. M. Shin (✉)
Electronics and Telecommunications Research Institute, 138 Gajeongno, Yuseong-gu, Daejeon 305-700, Republic of Korea
e-mail: hmshin@etri.re.kr
URL: http://heuristic.kaist.ac.kr/

C. Y. Lee
Department of Industrial & System Engineering, Korea Advanced Institute of Science and Technology, 335 Gwahangno, Yusung-gu, Daejeon 305-701, Republic of Korea
e-mail: chae@kaist.ac.kr

1 Introduction

According to the European Telecommunications Standards Institute (ETSI), WLAN-cellular integrated networks are classified into two types of coupling: loose coupling and tight coupling [1–3]. Tight coupling [4] shares information and resources between two wireless networks. In contrast, loose coupling can only provide seamless operation, because the networks operate independently of each other. In this case, a mobile station cannot properly share the overall resources of the two networks. If resources are to be used efficiently in loosely coupled wireless integrated networks, a resource management scheme is required that reflects the features of independent networks. Resource management across independent wireless networks is more difficult than in a single network. To address this difficulty, we consider the case of dual-mode mobile stations that move between two competitive wireless networks, namely a Universal Mobile Telecommunications System (UMTS) cellular network and a WLAN, in loosely coupled wireless integrated networks. We also address the problem of Call Admission Control (CAC) and rate allocation in wireless integrated networks.

Some related work has been done on problems pertaining to Radio Resource Management (RRM) in wireless integrated networks. This work has been based mainly on cellular networks and Wireless Local Area Networks (WLANs), which have an advantage with respect to coverage and capacity [5, 6]. Recently, related work has been extended to cover new wireless access technologies, such as Worldwide Interoperability for Microwave Access (WiMAX) and Wireless Metropolitan Area Networks (WMAN) [7, 8]. Approaches to RRM in WLAN-cellular integrated networks can be categorized as follows: WLAN-first schemes [9, 10], in which WLANs are always preferred by all services whenever the WLAN access is available; service specific schemes [11], in which the service is differentiated for QoS support; load balancing schemes [12], in which load balancing policies are designed to utilize the pooled resources of the network efficiently; and optimal joint session admission control schemes [13], in which an admission control scheme maximizes the overall network revenue with QoS constraints. In these studies, how to utilize the overall radio resource in loosely coupled wireless integrated networks optimally, subject to QoS constraints, has not been examined in detail. If the network providers are different, they will compete with each other to maximize their own revenue. In this situation, game theory can be used very profitably to address the relationship between the networks directly [8, 14]. In these game theory-based studies, given that each network is operated cooperatively, it was possible to maximize the overall and individual revenues of the providers. However, these game theory-based RRM schemes cannot be applied to systems in which the networks are operated independently and competitively, such as loosely coupled wireless integrated networks.

In this paper, we propose a scheme that exploits a mathematical model for the control of call admission and adopt a noncooperative game theory-based approach to address the rate allocation problem. The proposed schemes are based on certain features of networks and types of traffic. Generally, the price per packet (i.e., the unit price) in WLANs is less than in UMTS cellular networks. UMTS cellular networks basically guarantee some QoS requirements. In contrast, WLANs do not guarantee some QoS requirements. When price is the major consideration, users want to access WLANs, rather than UMTS cellular networks, whenever possible. However, when QoS is the major consideration, users want to access UMTS cellular networks, rather than WLANs, whenever possible. Real-time traffic is more sensitive to packet delay than nonreal-time traffic. In contrast, the nonreal-time traffic is more sensitive to packet loss than real-time traffic. Therefore, the criterion for call admission should differ according to the type of traffic. To reflect these features, we take packet-level QoS and the preferences of mobile stations into account in our proposed CAC scheme.

To develop an RRM scheme that reflects both these features of the network and the preferences of mobile stations, we present CAC and data rate allocation schemes for wireless integrated networks. The CAC scheme has two steps: (1) a network is selected according to the mobile station's preference, and (2) a decision is made as to whether or not to admit the call access request, on the basis of QoS requirements that are sensitive to the type of traffic. The data rate allocation scheme is based on noncooperative game theory, so as to maximize the revenue of the network providers while guaranteeing a certain level of QoS, which will vary according to user needs. In addition to providing an efficient approach to RRM, our second contribution is the mathematical modeling of packet-level QoS, with respect to the packet loss rate and packet delay time. Specifically, the model in UMTS cellular networks enables the loss rate and delay time to be predicted until the packets are transmitted. The model in WLANs allows the loss rate and delay time to be predicted until the packets have been transmitted successfully. In addition, we present a utility that can be used by mobile stations to decide which network they want to use. In contrast to previously presented schemes, when using our utility, a mobile station selects a network by comparing the cost and degree of satisfaction with QoS. The use of this utility provided important information about users' needs. With respect to the packet loss rate, packet delay time, and probability that a call will be blocked, we compare the proposed scheme with two other schemes. The comparison shows that the proposed scheme can guarantee a level of QoS according to user needs while maximizing the revenue of the two networks.

The remainder of this paper is organized as follows. In Sect. 2, the problem is described and a model for a wireless integrated network is defined. In Sect. 3, the packet-level QoS for each wireless network is predicted. In Sect. 4, the mobile stations' preferences are formulated and a CAC scheme is proposed to reflect the mobile stations' preferences and to respond sensitively to QoS requirements according to the type of traffic. In Sect. 5, a revenue-based rate allocation scheme is proposed that is based on two-person noncooperative game theory. In Sect. 6, a detailed description of the proposed scheme is presented, followed by a performance analysis and numerical results. Section 7 concludes.

2 Network model and features

The model under consideration is a loosely coupled architecture comprising a UMTS cellular network and a WLAN, as shown in Fig. 1. The wireless integrated network cannot control resource management, except for the roaming functions such as billing, authentication, and the mobility management mechanism between the two networks, because these networks are connected simply through gateways [15]. The mobile stations in this model are dual-mode stations; all

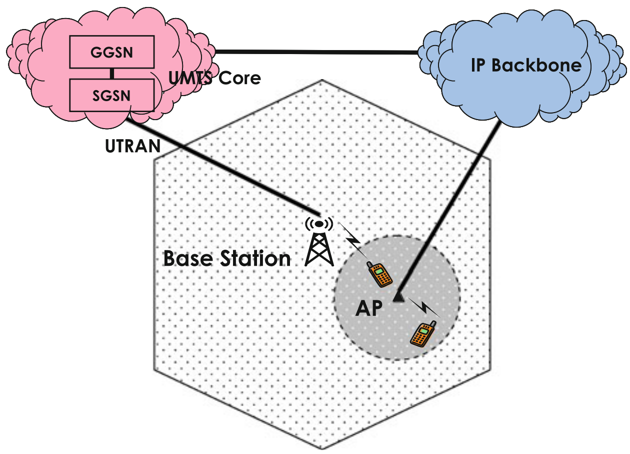

Fig. 1 The loosely coupled architecture in the UMTS cellular network/WLAN integrated network

of them have an interface card to access both networks. It is assumed that the entire coverage area of the wireless integrated network is covered by the UMTS cellular network, which serves low-bandwidth traffic. We assume that the WLAN is distributed randomly in the coverage area that is covered by the UMTS cellular network, and that the WLAN provides greater bandwidth service than does the UMTS cellular network. In the overlapping area, mobile stations may have more than one connection option. When a mobile station moves from the coverage area of one wireless network to that of the other at the boundary of the coverage area of an Access Point (AP), the mobile station can transfer from one network to the other.

Given that each network provider operates their own networks in the wireless integrated network individually, they will compete in order to maximize their revenue. To allocate the data rate optimally in this situation, a rate vector, which indicates a set of individual data rates that are allocated optimally to calls in each network, should be obtained independently. The problem of obtaining the rate vector independently can be addressed by using a two-person non-zero-sum noncooperative game model [16].

The two players of the game are two heterogeneous networks that compete with each other. Each network wants to increase the number of calls that its users make while allocating the rates in such a way as to maximize revenue. Note that these competing desires are related to the fact that the networks do not cooperate. Despite this competition, all of the available radio resources should be allocated effectively and fairly between the two wireless networks. This is a type of non-zero-sum game because the players do not have any conflicting objectives with respect to network utilization. Nonetheless, each network fixes its own prices according to the type of traffic. However, in many cases, there are some pairs of rate vectors such that one network allocates its rates without the cooperation of the other, and the network cannot improve its utility. Such points are called Nash equilibriums, in which no player has anything to gain by changing his or her own strategy unilaterally while the other players do not change.

3 Prediction for packet-level QoS

We here consider the delay time and packet loss rate, which multiple types of calls generate in a wireless network, as measures of the extent to which QoS is provided, and use them as decision metrics in the CAC. In a cellular system, packets sent and received by current users of the network are delayed whenever the signal to interference and noise ratio (SINR) is lower than the QoS requirements for each type of traffic. In a WLAN, packet delay occurs whenever calls contend in the network channels. As a result, packet losses occur if the packet delay time exceeds the QoS requirements for each type of traffic. It is assumed that the providers do not charge for lost packets, but only for packets that are transmitted successfully. Therefore, the more packets are lost, the less revenue is obtained.

Assume that there are $n_{k,j}$ calls for each type of traffic $k \in \{rt, nrt\}$, where rt and nrt denote real-time and non-real-time types of traffic, respectively, in network $j \in \{u, w\}$; Each call generates packets according to a Poisson process with rate $\delta_{k,j}\left(=\frac{r_{k,j}}{\beta_{k,j}} \times \varepsilon_j\right)$ per slot-time, $r_{k,j}$ is an allocated data rate for calls of traffic type k in network j, and $\beta_{k,j}$ is a packet size in bits. Each call transmits a data frame, which is divided into slot-time ε_j. One packet composed of $\beta_{k,j}$ bits is transmitted in one time slot. If the packet is not transmitted successfully, it can be retransmitted until the delay time exceeds the time limit, when the packet is dropped.

3.1 UMTS cellular network

In the UMTS cellular network, the proposed scheme is based on so-called complete sharing [17], which allows all mobile stations to use the available bandwidth equally. The UMTS cellular system consists of calls of multiple service classes. When a call is accepted, a data packet is generated immediately. The base station allocates different time slots for the accepted calls in the cell, such that a mobile station that has accepted a call can only transmit after it has been allocated a time slot and can transmit only for the period of the time slot.

The actual SINR $\gamma_k(i)$ for connection i, where $1 \le i \le n_u = \sum_{k\in\{rt,nrt\}} n_{k,u}$, is given by [18]

$$\gamma_k(i) = \frac{g_{i,k}P_{i,k}}{\theta_{h,k}\sum_{h=1,i\neq h}^{n_u} g_{h,k}P_{h,k} + N_{k,0}}, \quad \text{for} \quad k \in \{rt,\ nrt\} \tag{1}$$

where $P_{i,k}$ is the transmit power of the mobile station in connection i, $g_{i,k}$ denotes the path gain (=antenna gain/path loss) between the mobile station and the base station, $\theta_{h,k}$ denotes the reduction in interference due to signal processing (i.e., $\theta_{h,k} \approx 1/G_{i,k}$ for CDMA with processing gain $G_{i,k} = W/\upsilon_{i,k} r_{k,u}$ where W is the chip rate and $\upsilon_{i,k}$ is the activity factor for traffic class k), and $N_{k,0}$ denotes the thermal noise power received at the base station.

To guarantee the rate $r_{k,u}$, the actual SINR $\gamma_k(i)$ must be greater than or equal to the target SINR γ_k^*. If this constraint is not satisfied at the beginning of a time slot, some packets are not transmitted. To express this relationship between the SINR and packet transmission, the variable ζ_i, which indicates whether or not the connection i satisfied the target SINR γ_k^*, is introduced.

$$\zeta(i) = \begin{cases} 1, & \text{for } \gamma_k(i) \geq \gamma_k^* \\ 0, & \text{for } \gamma_k(i) < \gamma_k^* \end{cases} \tag{2}$$

Therefore, the packet throughput at a time slot is defined by $\kappa_{k,u} = \sum_{i=1}^{n_{k,u}} \zeta(i)$. For the mathematical derivation of packet-level QoS, with respect to the mean packet delay time and mean probability of packet loss, we assume that the packet throughput follows a Poisson distribution with the rate of $\kappa_{k,u}$. Based on the packet generation rate and throughput, the packet loss rate and delay time can be formulated in the UMTS cellular network. First, we can formulate the packet loss rate by characterizing the packet success probability.

Proposition 1 *For all packets generated from n_u calls in UMTS cellular network, the packet success probability $S_{k,u}$ can be expressed as*

$$S_{k,u} = 1 - \rho_{k,u} e^{-(\kappa_{k,u} - \delta_{k,u}) D_{k,u}}, \quad D_{k,u} \geq 0. \tag{3}$$

where $\rho_{k,u}$ is the packet load at a given time slot which can be denoted by $\rho_{k,u} = \delta_{k,u}/\kappa_{k,u}$; and $D_{k,u}$ is the maximum acceptable packet delay with no collision.

Proof Let $d_{k,u}$ be the transmission delay times in the queue, in which the accepted data messages wait before transmission in a base station. The delay time, denoted by $d_{k,u}$, is a random variable, distributed exponentially with mean $1/(\kappa_{k,u} - \delta_{k,u})$ since the packets are generated and depart according to Poisson process. Therefore, the probability that the packet is transmitted successfully is represented by $P(d_{k,u} \leq D_{k,u})$. If the packet delay of a call that is of traffic type k is smaller than $D_{k,u}$, the packet is considered to have been transmitted successfully; otherwise, it is considered to have failed. Packets that are not transmitted until the delay time meets the required delay bound $D_{k,u}$ are dropped. That is, the probability of successful transmission is expressed by the probability that the delay time is zero and that the delay time is less than or equal to $D_{k,u}$, provided that the delay time is greater than zero, as follows:

$$P(d_{k,u} \leq D_{k,u}) = P(d_{k,u} = 0) + P(d_{k,u} > 0) P(d_{k,u} \leq D_{k,u} | d_{k,u} > 0). \tag{4}$$

$P(d_{k,u} = 0)$ in (4) is the probability that the delay time is zero, $P(d_{k,u} = 0) = 1 - \rho_{k,u}$. $P(d_{k,u} > 0)$ in (4) is the probability that the delay time is greater than zero, $P(d_{k,u} > 0) = 1 - P(d_{k,u} = 0) = \rho_{k,u}$. $P(d_{k,u} \leq D_{k,u} | d_{k,u} > 0)$ in (4) is the conditional probability of $d_{k,u} \leq D_{k,u}$ given $d_{k,u} > 0$. It is equivalent to $1 - P(d_{k,u} > D_{k,u} | d_{k,u} > 0)$. Due to the fact that the exponential distribution is completely independent from the previous events in each subsequent event, $P(d_{k,u} > D_{k,u} | d_{k,u} > 0)$ is equivalent to $P(d_{k,u} > D_{k,u})$. Then $P(d_{k,u} > D_{k,u}) = e^{-(\kappa_{k,u} - \delta_{k,u}) D_{k,u}}$, and so $P(d_{k,u} > D_{k,u} | d_{k,u} > 0) = e^{-(\kappa_{k,u} - \delta_{k,u}) D_{k,u}}$. Therefore, the probability that the packet will be transmitted successfully $P(d_{k,u} \leq D_{k,u})$ is expressed by $P(d_{k,u} \leq D_{k,u}) = 1 - \rho_{k,u} + \rho_{k,u}[1 - e^{-(\kappa_{k,u} - \delta_{k,u}) D_{k,u}}] = 1 - \rho_{k,u} e^{-(\kappa_{k,u} - \delta_{k,u}) D_{k,u}}, \quad D_{k,u} \geq 0$. This completes the proof. □

Therefore, the probability that a particular packet will be lost for traffic type k in the UMTS cellular network is denoted by $L_{k,u} = 1 - S_{k,u}$. From [19], $\bar{D}_{k,u} = E[d_{k,u}] = \delta_{k,u}/\kappa_{k,u}(\kappa_{k,u} - \delta_{k,u})$ is the mean packet delay time for calls of traffic type k in the UMTS cellular network. That is, the mean packet delay time for calls of traffic type k is defined in terms of the packet generation rate and throughput.

3.2 WLAN

Consider a single WLAN under the assumption of ideal channel conditions. The distributed coordination function (DCF) access method of the IEEE 802.11 standard [20] is based on the carrier-sense multiple-access with collision-avoidance (CSMA/CA) protocol. A station that has a packet to transmit senses the channel and, if the channel remains free for a distributed interframe space (DIFS) time, it transmits the packet. If the station determines that the channel is busy, it waits until the channel becomes idle for a DIFS time, after which it initiates a backoff process.

The IEEE 802.11 MAC performs an exponential backoff if a station is involved in a collision. When there is a collision, each station chooses a discrete random variable, $f(0, \mathrm{CW}_m - 1)$, from a uniform distribution in the range $\{0, 1, \ldots, \mathrm{CW}_m - 1\}$, where CW_m is the contention window and m is the number of collisions, and initializes its backoff time counter with this number. That is, the number of collisions, m, is the number of trials needed for the first success. Therefore, the number of collisions, m, is a random variable distributed geometrically as follows:

$$p(m) = \left(\frac{n_w - 1}{n_w}\right)^m \left(\frac{1}{n_w}\right), \quad m = 0, 1, \ldots; \quad 0 < \frac{1}{n_w} < 1. \tag{5}$$

When this mechanism is used, a packet sent by a mobile station is lost when the number of collisions is more than the retry limit M. Therefore, a packet must encounter fewer than M collisions if it is to be transmitted successfully. This condition can be expressed for each traffic class by the following:

$$S_{k,w} = p(m \le M) = \sum_{m=0}^{M} \left(\frac{n_w - 1}{n_w}\right)^m \left(\frac{1}{n_w}\right), \tag{6}$$

where $S_{k,w}$ is the probability that a packet that is part of a call of traffic type k will be transmitted successfully in the WLAN, and $n_w = \sum_{k \in \{rt,nrt\}} n_{k,w}$ is the number of calls in the WLAN. Therefore, the probability that a packet that is part of a call of traffic type k will be lost in the WLAN is denoted by $L_{k,w} = 1 - S_{k,w}$.

The above means that a packet is transmitted successfully if the number of retransmissions is fewer than the retry limit. In this case, the backoff delay, the length of packets that are transmitted successfully, and the length of packets that have encountered collisions are considered, as shown in Fig. 2. First, the backoff delay for each packet is the time taken, taking into account the backoff time and the number of collisions that have occurred. Then, the total backoff time until a packet is transmitted successfully is expressed as follows:

$$\begin{aligned} T_{\text{backoff}} = \varepsilon_w\{&p(0)f(0, \text{CW}_0 - 1) \\ &+ \cdots + p(m)f(0, \text{CW}_m - 1)\}, \end{aligned} \tag{7}$$

where $\text{CW}_m = \min(2^m \text{CW}_{\min}, \text{CW}_{\max})$ and $m \le M$. A packet that has encountered a collision is retransmitted after a random number of time slots. At the first transmission attempt, CW_m is set to a value $\text{CW}_{\min}$. It is then doubled after each unsuccessful transmission, up to a maximum value of $\text{CW}_{\max}$. The $\text{CW}_{\min}$ and $\text{CW}_{\max}$ values are fixed. If there are collisions, they occur within the length of time that is assigned to a packet for transmission. The packet delay, according to the RTS frame length and DIFS time, is expressed as $m \cdot T_{\text{collision}}$, where $T_{\text{collision}}$ is the sum of the RTS frame length and DIFS time. Lastly, the length of the successful packet header and payload T_{success} are considered. When the RTS/CTS option is used, successful packet transmissions are preceded by an RTS/CTS exchange, while collisions occur between RTS frames instead of data packets. Therefore, the delay time (d_k) until a packet is successfully transmitted is a summation of the backoff delay (T_{backoff}), the length of successfully transmitted packets (T_{success}), and the length of collided packets ($T_{\text{collision}}$), as follows:

$$\left.\begin{aligned} d_{k,w} &= T_{\text{backoff}} + m \cdot T_{\text{collision}} + T_{\text{success}} \\ T_{\text{success}} &= T_{\text{RTS}} + T_{\text{SIFS}} + T_{\text{CTS}} + T_{\text{DIFS}} + T_{\text{PHY}} \\ &\quad + T_{\text{MAC}} + T_{\text{PAYLOAD}} + T_{\text{SIFS}} + T_{\text{ACK}} + T_{\text{DIFS}} \\ T_{\text{collision}} &= T_{\text{RTS}} + T_{\text{DIFS}} \end{aligned}\right\}. \tag{8}$$

The packet delay time $d_{k,w}$ is also a random variable with the same distribution because the delay time is a linear function of the constant and random variable m. In (8), only T_{backoff} and $T_{\text{collision}}$ among the elements that comprise the delay time are related to the random variable m. Therefore, the delay time is affected by random variable m. Based on those properties, it is possible to formulate the mean packet delay time.

Proposition 2 *For all packets generated from n_w calls in WLAN, the mean packet delay time, denoted by $\bar{D}_{k,w}$, can be expressed by*

$$\begin{aligned} \bar{D}_{k,w} = \; &\varepsilon_w\{p(0)f(0, \text{CW}_0 - 1) + \cdots + p(n_w) \\ &\times f(0, \text{CW}_{n_w} - 1)\} + n_w \cdot T_{\text{collision}} + T_{\text{success}}. \end{aligned} \tag{9}$$

Proof The mean packet delay time $\bar{D}_{k,w}$ can be represented by

$$\bar{D}_{k,w} = E[d_{k,w}], \tag{10}$$

where $E[d_{k,w}]$ is an expected value of the packet delay time. In (10) above, since the backoff delay time and collision delay time are affected by the random variable m, we need to obtain the expected value of the random variable m first. The expected value of the random variable m is as follows:

$$\begin{aligned} E[m] &= \sum_{m=0}^{\infty} m \left(\frac{n_w - 1}{n_w}\right)^m \left(\frac{1}{n_w}\right) = \left(\frac{1}{n_w}\right) \sum_{m=0}^{\infty} m \left(\frac{n_w - 1}{n_w}\right)^m \\ &= \left(\frac{1}{n_w}\right) \frac{1}{\left\{1 - \left(\frac{n_w - 1}{n_w}\right)\right\}^2} \\ &= n_w. \end{aligned} \tag{11}$$

Therefore, the expected number of collisions, denoted by $E[m]$, is applied to the backoff and collision delay time in (8) as follows:

$$\begin{aligned} E[d_{k,w}] &= E[T_{\text{backoff}} + m \cdot T_{\text{collision}} + T_{\text{success}}] = E[T_{\text{backoff}}] \\ &\quad + E[m \cdot T_{\text{collision}}] + E[T_{\text{success}}] \\ &= E[\varepsilon_w\{p(0)f(0, \text{CW}_0 - 1) + \cdots \\ &\quad + p(E[m])f(0, \text{CW}_{E[m]} - 1)\}] \\ &\quad + E[m] \cdot E[T_{\text{collision}}] + E[T_{\text{success}}] \\ &= \varepsilon_w\{p(0)f(0, \text{CW}_0 - 1) + \cdots + p(n_w) \\ &\qquad f(0, \text{CW}_{n_w} - 1)\} + n_w \cdot T_{\text{collision}} + T_{\text{success}}. \end{aligned} \tag{12}$$

This completes the proof. □

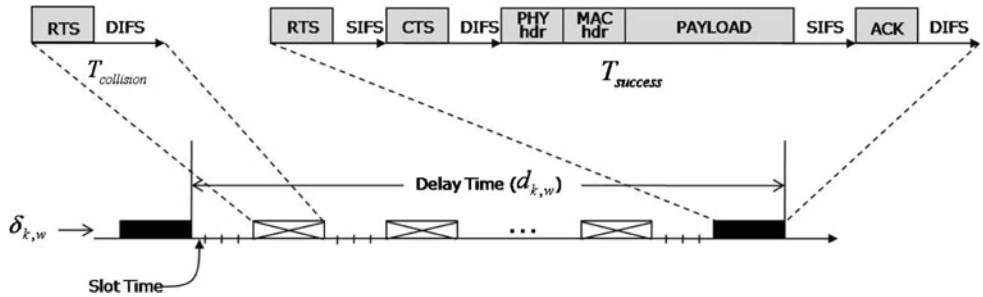

Fig. 2 The delay time between successful packet transmissions

4 Call admission control

4.1 Traffic model

We consider an analytic traffic model where two types of traffic, such as real-time and non-real-time, access to the considering network model. In the overlapped area, where UMTS cellular network's cell overlays the WLAN's zone, the arrivals of both new calls and vertical handoff calls from UMTS cellular network are Poisson distributed with arrival rate λ_k. That is, if θ percent of arrival rate λ_k is new calls, then the rest $(100-\theta)$ percent of arrival rate λ_k is vertical handoff calls. Call duration time for traffic class k is exponentially distributed with the mean $1/\mu_k$. The above exponential call duration time has also been shown to be valid for wide range of systems [21, 22]. Of course, the call arrival rates and duration times in each network can be different. For simplicity performance evaluation, we don't divide these between two networks.

4.2 Mobile station's preference

We consider an active handoff, such that a mobile station acquires some information or signal strength from a base station or access point. In this case, when a mobile station sends a request signal to the neighborhood, base stations and access points send a response signal. In the active handoff, we assume that a mobile station acquires some information from the response packet. Therefore, a mobile station has some information upon which to base its preference, according to the type of traffic. For each network, a mobile station has a gain, $g_{k,j} \geq 0$, which represents the degree of QoS received by the mobile station when compared with the service price per packet, but has no gain for failed packets. The network providers earn revenue equal to the unit price. That is, the more delay a packet experiences, the less gain the mobile station obtains. This relationship is given by

$$g_{k,j} = \begin{cases} c_{k,j}, & d_{k,j} = 0 \\ c_{k,j}/d_{k,j}, & 0 < d_{k,j} \leq D_{k,j} \end{cases}, \tag{13}$$

where $c_{k,j}$ is the price per packet for each call of traffic type k in network j. To equalize the measure between gain and price, the unit price can be transformed to the unit price per time slot because a packet is transmitted during one time slot. The UMTS cellular packet radio network provides services with unit prices $c_{k,u}$ for each traffic type k. Meanwhile, the WLAN provides services with unit prices $c_{k,w}$ for each traffic type k. Note that the unit price in the cellular network is greater than in the WLAN, i.e. $c_{k,u} > c_{k,w}$.

If a call of traffic type k is admitted in network j when the number of calls is n_j for each network, the expected utility of the mobile station, which represents the relationship between a mobile station's payment and the QoS that is provided during the call, can be calculated as follows:

$$U(k,j) = \delta_{k,j} S_{k,j} (\bar{g}_{k,j} - c_{k,j}) / \varepsilon_j \mu_k, \tag{14}$$

where the expected gain $\bar{g}_{k,j}$ is defined as

$$\bar{g}_{k,j} = \begin{cases} c_{k,j}, & E[d_{k,j}] = 0 \\ c_{k,j}/E[d_{k,j}], & 0 < E[d_{k,j}] \leq D_{k,j}. \end{cases} \tag{15}$$

Given unit prices, the mobile station has a preference for the network having a larger utility value. That is, if the mobile station's utility in network j is higher than in network j', $U(k,j) > U(k,j')$, the mobile station will first access network j. However, if the mobile station's utility in network j is equal to its utility in network j', $U(k,j) = U(k,j')$, then the mobile station will randomly access network j and j' because it has no preference for one or the other.

4.3 Traffic-type-sensitive CAC

If a mobile station's network preference has been decided in the above step, the CAC scheme proceeds to the next decision-making step, where the network admits or rejects the call access. Given the status of each network, the network must decide whether it will admit or reject the applying calls. At this time, the current level of QoS provided is considered.

The packet-level QoS is to guarantee the host-to-host QoS. In the network architecture, the network layer is responsible for guaranteeing and controlling the host-to-host transmission. Also the handoff is dealt with in the area of IP mobility protocol, and done in network layer. Our proposed scheme is to allocate the data rate for handoff call as well as new call. Therefore, the packet-level QoS should be calculated and guaranteed in network layer when the networks decide whether they admit a new/handoff call or not, and how much data rates do they allocate to those calls, so that in this paper it is possible to apply it to the CAC in the network layer. Therefore, packet delay and packet loss are used to determine the level of QoS provided. That is, the mean packet delay time will be applied to the real-time traffic types, and the probability that a particular packet will be lost will be applied to the non-real-time traffic types.

This CAC scheme provides mobile stations with guaranteed QoS at the packet level and call level by applying the mobile stations' preferences. Therefore, the CAC scheme follows he procedure as shown in Fig. 3.

5 Optimal rate allocation

5.1 Network revenue

To allocate the data rate fairly among networks, game theory is applied to derive a solution that results in reciprocal revenue between the UMTS cellular network and the WLAN. The expected network revenue is obtained under the assumption that the data rates of the corresponding calls are fixed for the duration of the calls if the data rate is allocated to the calls. That is, these packets are generated continuously for the duration of the calls; hence, the expected total revenue for each network can be obtained as follows.

The network revenue is an expected value for all ongoing calls and a new call at a particular decision time. Although call duration is continuous, it can be divided into time slots. Therefore, the revenue from a call of traffic type k that is gathered while the call is in progress is expressed as follows:

$$\mathrm{V}_{k,j} = \frac{c_{k,j} S_{k,j}}{\beta_{k,j}} r_{k,j}. \tag{16}$$

From (16) above, the total expected revenue from all calls of all types of traffic can be expressed by the following:

$$\mathrm{V}_j = \sum_{k=1}^{K} n_{k,j} \mathrm{V}_{k,j}. \tag{17}$$

This function is the expected revenue in a network from the packets that are generated during calls of all types of traffic. It is shown in (17) that the total expected revenue of network j is affected by the number of calls for traffic type k, the price per packet, packet success probability, and packet size as well as data rate. Thus, the revenue can be represented by a logarithm function [23] as follows:

$$\mathrm{V}_j = \log(\alpha_j r_j), \tag{18}$$

where α_j is a coefficient associated with the parameters such as $n_{k,j}$, $c_{k,j}$, $S_{k,j}$, and $\beta_{k,j}$; and $r_j = \sum_{k \in \{rt,nrt\}} r_{k,j}$. Next, a noncooperative game approach will be applied to study the competition between the UMTS cellular network and the WLAN.

5.2 Application of noncooperative game theory

5.2.1 Formulation of the noncooperative rate allocation game (NRAG)

There is no bilateral cooperation regarding the allocation of data rates between two loosely integrated networks. Rather, the two networks will compete with each other to maximize their revenues. Each network must allocate at least one data rate to meet the QoS requirements for each traffic type k. In this competitive environment, each network tries to maximize its own revenue while satisfying the QoS requirements for each traffic type k. There is thus a problem about how to allocate data rates. To solve this rate allocation problem, a two-person non-zero-sum noncooperative game model can be constructed for the available networks in the overlapping area.

The players in this game are available networks in the overlapping area. When a call of traffic type k is arriving, the strategy of each player is to allocate data rates so as to guarantee the QoS requirements for the call. R_j is the strategy space is constrained by $R_j = [0,\ C_j]$, where C_j is the total bandwidth capacity of network j. Each network allocates a data rate r_j such that $r_j \in R_j$. That is, a data rate r_j belongs to the continuous bound $[0,\ C_j]$.

The payoff function of a network can be represented as a profit and loss. The profit of a network is obtained by subtracting the revenue generated by ongoing calls from the revenue generated due to allocating the data rate to a new call in addition to ongoing calls. The loss of a network can be represented as a kind of chance loss. That is, the chance loss is a loss expected due to sharing the allocation of data rate with the other networks, although a network can obtain the whole revenue if the data rate is fully allocated only by the network. Therefore, the payoff of network j in allocating the data rate r_j to all calls in network j can be expressed as

Fig. 3 Pseudo-code for "preference + QoS-sensitive" CAC scheme

1: **Input**: $n_1, \cdots, n_u : n_1, \cdots, n_w$
2: $n_j \leftarrow n_j + 1$
3: Calculating $U(k, j)$ for each network j
4: **if** $U(k,u) > U(k,w)$ **then**
5: **while** $k = rt$
6: **if** $\overline{D}_{k,w} \le D_k^*$ **then**
7: Accept the request
8: $n_w \leftarrow n_w + 1$
9: exit
10: **else**
11: **if** The request has already tried to WLAN **then**
12: Reject the request
13: exit
14: **else**
15: Go to line 28
16: **endwhile**
17: **while** $k = nrt$
18: **if** $L_{k,w} \le L_k^*$ **then**
19: Accept the request
20: $n_w \leftarrow n_w + 1$
21: exit
22: **else**
22: **if** The request has already tried to WLAN **then**
23: Reject the request
24: exit
25: **else**
26: Go to line 28
27: **endwhile**
28: **else if** $U(k,u) < U(k,w)$ **then**
29: **while** $k = rt$
30: **if** $\overline{D}_{k,u} \le D_k^*$ **then**
31: Accept the request
32: $n_u \leftarrow n_u + 1$
33: exit
34: **else**
35: **if** The request has already tried to UMTS cellular network **then**
36: Reject the request
37: **else**
38: Go to line 4
39: **endwhile**
40: **while** $k = nrt$
41: **if** $L_{k,u} \le L_k^*$ **then**
42: Accept the request
43: $n_u \leftarrow n_u + 1$
44: exit
45: **else**
46: **if** The request has already tried to UMTS cellular network **then**
47: Reject the request
48: **else**
49: Go to line 4
50: **endwhile**
51: **else** $U(k,u) = U(k,w)$ **then**
52: Radom select a network j
53: **while** $k = rt$
54: **if** $\overline{D}_{k,j} \le D_k^*$ **then**
55: Accept the request
56: $n_j \leftarrow n_j + 1$
57: exit
58: **else**
59: Reject the request
60: exit
61: **endwhile**
62: **while** $k = nrt$
63: **if** $L_{k,j} \le L_k^*$ **then**
64: Accept the request
65: $n_j \leftarrow n_j + 1$
66: exit
67: **else**
68: Reject the request
69: exit
70: **endwhile**

$$\Phi_j(r_j, r_{-j}) = \log(\alpha_j r_j - \alpha_j' r_j') - \{\log(\alpha_j'' r_j'') - \log(\alpha_j r_j + \alpha_{-j} r_{-j})\}, \quad \text{for } j \in \{u, w\}, \tag{19}$$

where $\alpha_j(\alpha_{-j})$ is a coefficient associated with the data rate $r_j(r_{-j})$ which is allocated to ongoing calls together with a new call in network $j(-j)$; α_j' is a coefficient associated with the data rate r_j' which is allocated to ongoing calls in network j; and α_j'' is a coefficient associated with the data rate r_j'' which ongoing calls together with a new call in network j and $-j$ are allocated only by network j.

The network achieves r_j^* such that (19) is maximized. That is, the NRAG is expressed as

$$\max_{r_j \in R_j} \Phi_j(r_j, r_{-j}), \quad \text{for } j \in \{u, w\}. \tag{20}$$

The objective of the NRAG problem is to find out the equilibrium points, at which both networks cannot change the rates without cooperation to obtain more utility.

5.2.2 Nash equilibrium in the NRAG

In order to obtain a solution in such a NRAG problem, the definition of Nash equilibrium [24] is first introduced.

Definition 1 A rate vector $\mathbf{r} = (r_j, r_{-j})$ is a Nash equilibrium of the non-cooperative rate allocation game if, for every $j \in \{u, w\}$, $\Phi_j(r_j, r_{-j}) \geq \Phi_j(\hat{r}_j, r_{-j})$ for all $\hat{r}_j \in R_j$.

In a Nash equilibrium, none of the players can unilaterally change its strategy to increase its utility, as shown in Fig. 4.

5.2.3 Existence and uniqueness of Nash equilibrium in the NRAG

The Nash equilibrium does not necessarily exist. First, we investigate the existence of an equilibrium in the NRAG.

Theorem 1 *A Nash equilibrium exists in the NRAG.*

Proof The Nash equilibrium exists only when these two conditions are satisfied [25, 26].

1. r_j is a nonempty, convex, and compact subset of some Euclidean space $\Re^2$.
2. $\Phi_j(r_j, r_{-j})$ is continuous in $\mathbf{r}$ and qusi-concave in r_j.

Each network has a strategy space that is defined by a minimum rate, a maximum rate, and all the rate values in between. We also assume the maximum rate is larger than or equal to the minimum rate. Thus, the strategy space R_j of each network is a compact, convex set with minimum and maximum data rate constraints denoted by 0 and C_j, respectively.

And we can know that $\Phi_j(r_j, r_{-j})$ in (19) is a continuous function. Then the second-order partial derivative of $\Phi_j(r_j, r_{-j})$ with respect to r_j is given by

$$\frac{\partial^2 \Phi_j(r_j, r_{-j})}{\partial r_j^2} \leq 0, \quad \text{for } j \in \{u, w\}. \tag{21}$$

In (21) above, the equality is valid when the probability that a packet will be transmitted successfully is zero. Therefore, it meets the property of quasi-concavity. Since the NRAG satisfies the above two conditions, we can know that our NRAG problem should have at least one Nash equilibrium. This completes the proof. □

And then, we turn to prove the uniqueness of Nash equilibrium.

Theorem 2 *The Nash equilibrium of the NRAG is unique.*

Proof To prove the uniqueness of the Nash equilibrium, we obtain the best response function of each network j by differentiating the payoff function in (19) with respect to r_j. From (19), the first-order partial derivative of payoff function is given by

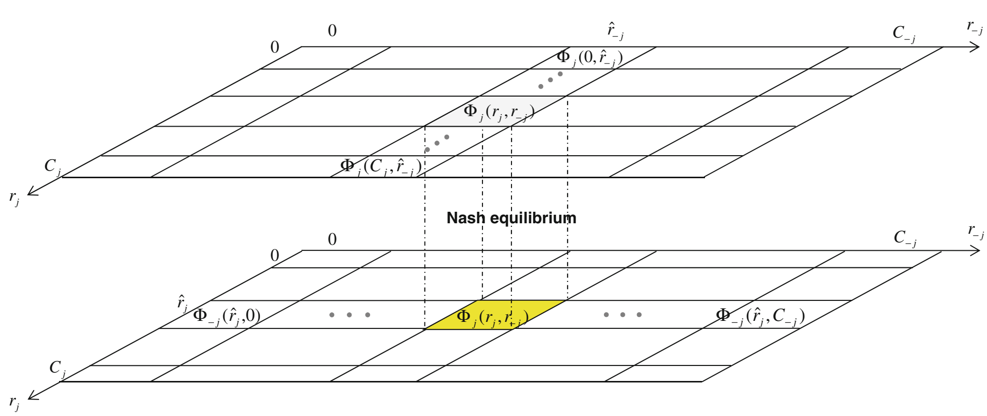

Fig. 4 Optimal rate allocation scheme based on non-cooperative game theory

$$\frac{\partial \Phi_j(r_j, r_{-j})}{\partial r_j} = \frac{\partial \log(\alpha_j r_j - \alpha'_j r'_j)}{\partial r_j} - \frac{\partial\{\log \alpha''_j r''_j - \log(\alpha_j r_j + \alpha_{-j} r_{-j})\}}{\partial r_j}, \quad (22)$$

Then, the best response function for network j can be expressed as follows:

$$r^*_j = \left(\frac{\alpha'_j}{2\alpha_j}\right) r'_j - \left(\frac{\alpha_{-j}}{2\alpha_j}\right) r_{-j}, \quad \text{for}\, j \in \{u, w\}, \quad (23)$$

In (23) above, there is a unique intersection if and only if the slopes of best response function for each network are not equal [27, 28]. Since the slopes are unequal, there is a unique intersection which is the equilibrium. This completes the proof. □

In (23), the slopes of the best response function are dependent on α_j and α_{-j}. Since the packet size and call duration time are established under the same condition for each network, the slopes are affected by the number of calls for traffic type k, the price per packet, and packet success probability. The change of Nash equilibrium according to the change of parameters, such as the price per packet and packet success probability, is shown as Fig. 5(a) and (b). Figure 5(a) shows the best response functions for the UMTS cellular network and the WLAN. If the data rates allocated to ongoing calls in the UMTS cellular network and the WLAN are 150 and 200 kbps, respectively, there is a Nash equilibrium point in each network. If the price per packet in one network is increased, the equilibrium point is shifted to the right-hand side, and vice versa. Therefore, the more the price per packet is, the more the allocated data rate in one network is. Under the same condition of data rate allocated to ongoing calls as above, Fig. 5(b) shows the variation of best response functions for the UMTS cellular network and the WLAN under different probability that a packet is transmitted successfully in the overlapping area. If the probability that a packet is transmitted successfully in one network is increased, the equilibrium point is shifted to the right-hand side, and vice versa. Therefore, the more the probability that a packet is transmitted successfully is, the more the allocated data rate in one network is. For the rest, the number of calls shows the same variation as above two components. Only the packet size shows the completely opposite pattern.

6 Performance evaluation

We evaluated the proposed scheme by comparing various performance metrics, including the probability that new calls will be blocked, the probability that vertical handoff calls will be dropped, and revenue. We now report on the evaluation. The numerical experiments were done using the simulator ns-2.29, which was developed by the National Institute of Standards and Technology (NIST) [29]. We study the performance of QoS-sensitive admission control (QoSAC) scheme that considers the noncooperative game theory-based rate allocation (NRAG). We also study the effect of the QoSAC scheme by comparing it with the QoS support [11] and optimal joint session admission control (JSAC) [13] schemes.

In the simulations, the evaluated wireless integrated networks consisted of one UMTS cellular network and one WLAN, where the UMTS cell coverage fully covers that of the WLAN. The bandwidths of the UMTS cellular network and the WLAN were 384 kbps and 11 Mbps, respectively. The traffic in the integrated network, including speech, video, world wide web, and e-mail, can be categorized into six classes as shown in [30]. The simulation considered two types of traffic: real-time (speech) and non-real-time (recorded video). The 95% confidence intervals of the simulation results in the following figures originate from 10 independent runs. At each point in those figures, the width of the confidence interval for the population mean is $2z_{0.05/2}\sigma/\sqrt{x}$, and σ and x are the standard deviation and the total number of calls generated during the simulation, respectively. For each run, the simulated time was 1,000 units of the mean connection holding time. The initial 100 time units were considered to be the transient period. Performance samples for this period were discarded.

According to [31, 32], the average delay time of real-time traffic should be less than the delay tolerance $D^*_{rt} = 150$ ms. Further, probability that a particular packet will be lost $L^*_{rt} = 0.03$ for real-time traffic is acceptable with a data rate of 4–64 kbps. In contrast to real-time traffic, the average delay time of non-real-time traffic should be less than the delay tolerance $D^*_{nrt} = 10s$, while probability that a particular packet will be lost $L^*_{nrt} = 0.01$ for non-real-time traffic is acceptable within 28.8–384 kbps.

We assumed that the desired packet rate for each type of traffic is decided by the applications. The data rates for each type of traffic in the UMTS cellular network and the WLAN belong to the continuous bound [4, 64 kbps] and [28.8, 384 kbps], respectively. The numerical values for the WLAN physical layer parameters are shown in Table 1. The PHY and MAC header sizes are 24 and 30 bytes, respectively. The numerical values for the WCDMA physical layer parameters are shown in Table 2. When data is transmitted in the network, application data that is generated with a payload size that is fixed in the MAC frame is transmitted. If the size of a generated packet exceeds that of the payload, the application splits the packet into smaller

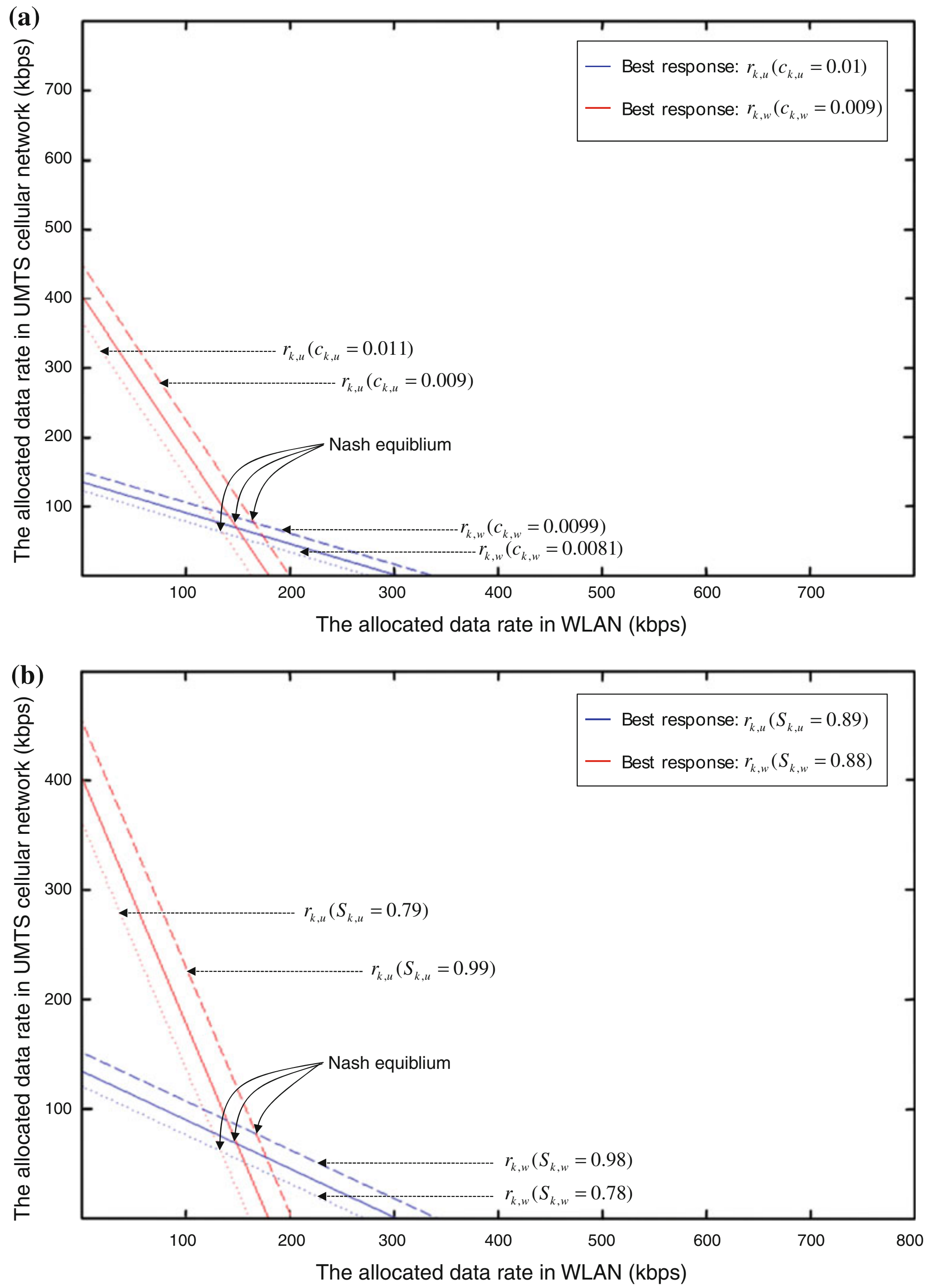

Fig. 5 Sensitivity for (**a**) the cost and (**b**) the probability that a packet is transmitted successfully in the best response functions

pieces. Large packets are split into smaller ones in order to satisfy the constraint on the size of the payload that is fixed in the MAC frame. Therefore, it is reasonable that each type of traffic generates packets of the same size. And the path loss between the mobile station and the base station, $PL_{i,k}$, is given by the values obtained from the following equation in [33].

$$\begin{aligned} PL_{i,rt} = \; & 18.0[\text{dBm}] - \{5.0[\text{dB}] - 10\log(3840/r_{rt,u})[\text{dB}] \\ & + (-100.2)[\text{dBm}]\} + 18.0[\text{dBi}] - 2.0[\text{dB}] \\ & - 4.0[\text{dB}] \\ PL_{i,nrt} = \; & 26.0[\text{dBm}] - \{1.5[\text{dB}] - 10\log(3840/r_{nrt,u})[\text{dB}] \\ & + (-100.2)[\text{dBm}]\} + 18.0[\text{dBi}] - 2.0[\text{dB}] \\ & - 4.0[\text{dB}] \end{aligned}$$

Table 1 WLAN parameters

Parameter	Notation	Value
Bandwidth capacity	C_w	11 Mbps
Slot-time	ε_w	10 µs
Time required to transmit ACK	T_{ACK}	203 µs
Short interframe space (SIFS) time	T_{SIFS}	10 µs
Distributed interframe space (DIFS) time	T_{DIFS}	50 µs
Packet size	$\beta_{k,w}$	512 byte
Minimum and maximum of contention window	$CW_{\min}$, $CW_{\max}$	31, 1,023

Table 2 UMTS cellular network parameters

Parameter	Notation	Value
Bandwidth capacity	C_u	384 kbps
Slot-time	ε_u	2,560 chips (=666 µs)
SINR	γ_{rt}^*, γ_{nrt}^*	3, 1 dB
Transmit power of mobile station	$P_{i,rt}$, $P_{i,nrt}$	125 mW(21 dBm), 250 mW (24 dBm)
Chip rate	W	3.84 Mcps
Receiver noise power	$N_{k,0}$	−103.2 dBm
Packet size	$\beta_{k,u}$	512 byte
Activity factor	$\upsilon_{i,rt}$, $\upsilon_{i,nrt}$	0.67, 1.00
Antenna gain	–	2.0 dBi

6.1 Optimal establishment of the unit price ratio

Figure 6 shows the variations in the probability that new calls will be blocked against the ratio of unit price between the two networks. If 50, 60, and 70% of the total number of calls that arrive carry real-time data, $\lambda_{rt} = 0.05$ ($\lambda_{nrt} = 0.05$), $\lambda_{rt} = 0.06$ ($\lambda_{nrt} = 0.04$), and $\lambda_{rt} = 0.07$ ($\lambda_{nrt} = 0.03$), respectively. In this example, $1/\mu_{rt} = 0.005(1/\mu_{nrt} = 0.005)$. As expected, it was found that the probability that new calls will be blocked in the WLAN falls as the difference between the unit prices in each network decreases. However, the probability that new calls will be blocked in the UMTS network increases. Almost all of the mobile stations prefer the WLAN when the price per packet in the WLAN is considerably less than that in the UMTS cellular network. The less the difference in the price per packet between the two networks, the greater is the preference for the UMTS cellular network. The two lines intersect at unit price ratios of 0.66, 0.58, and 0.50 for 50, 60, and 70% of real-time calls in both networks, respectively. At these intersections, the probability that new calls will be blocked is the same in the two networks. That is, if the price per packet is set at these points, equal network performance is achieved because the traffic load is balanced between the two networks.

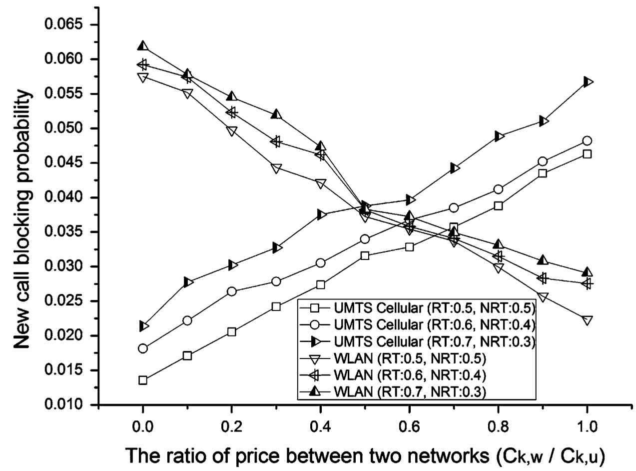

Fig. 6 The probability that new calls will be blocked against the unit price ratio between the UMTS cellular network and the WLAN

6.2 QoS performance in the QoSAC

Figure 7(a) and (b) show the performance of our QoSAC scheme and the other schemes in terms of mean packet delay time in the UMTS cellular network and WLAN, respectively. The parameters in this example are set as follows: $1/\mu_{rt} = 0.005(1/\mu_{nrt} = 0.005)$, $c_{rt,u} = \$0.01$ ($c_{nrt,u} = \$0.006$), and $c_{rt,w} = \$0.009$ ($c_{nrt,w} = \$0.002$). Under the same call duration time, the call arrival rate represents the traffic load of a network. That is, a high arrival rate indicates a high traffic load. As seen in Fig. 7(a) and (b), the performance improvement for real-time and nonreal-time traffics in our scheme is obvious. Especially, the improvement increases with higher arrival rate. When the call arrival rate increases, the mean packet delay times in the QoSAC scheme is always lower than the delay tolerance, because only this scheme consider the constraints on packet-level QoS in both networks. This shows that the QoSAC scheme is even more efficient with respect to packet delay time than the other schemes, because the QoSAC scheme does not only admit the call but also allocate the data rates if it is predicted not to satisfy the packet delay time.

Figure 8(a) and (b) show the variations in the probability that a particular packet will be lost against the call arrival rate in the overlapping area. In line with the aims of the schemes, the probability that a particular packet will be lost also satisfies the tolerance for packet loss in both the QoSAC and JSAC schemes, because only these schemes consider the constraints on packet-level QoS in both networks. Moreover, the packet loss probabilities of real-time and nonreal-time traffics in QoSAC scheme are slightly lower than that of the JSAC scheme. This is also because the QoSAC scheme does not only admit the call but also allocate the data rates if it is predicted not to satisfy the packet loss probabilities. Therefore, it is clear that the

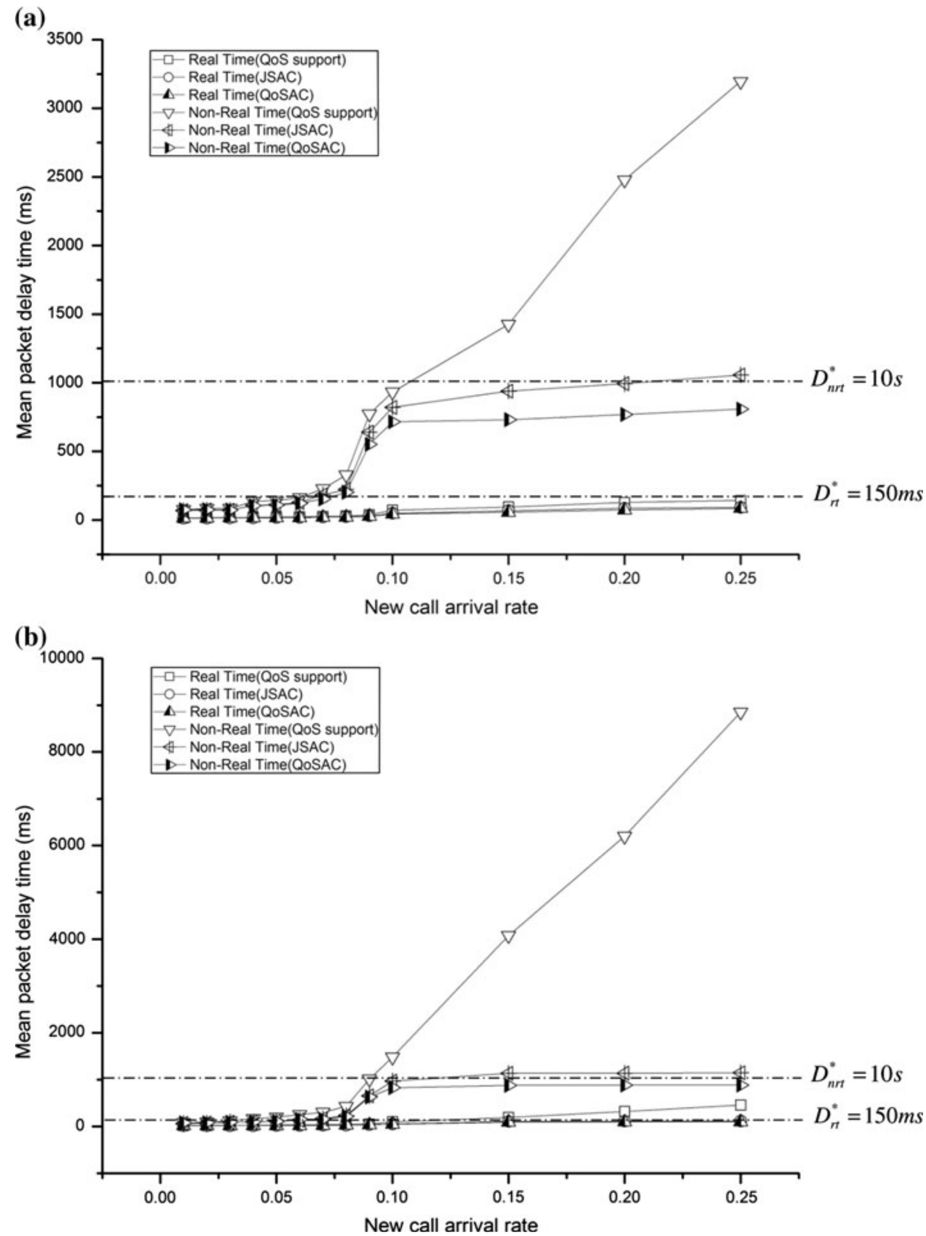

Fig. 7 Mean packet delay time in (**a**) the UMTS cellular network and (**b**) the WLAN

QoSAC scheme is more efficient with respect to packet-level QoS.

Figure 9(a) and (b) show the variations in the probability that new calls will be blocked and the probability that vertical handoff calls will be dropped against the rates at which new calls arrive. Let the handoff calls be equal to 50% ($\theta = 50$) of these incoming calls. In Fig. 9(a), the probability that new calls will be blocked at the new call arrival rate 0.25 is 0.27, 0.15, and 0.11 in the QoS support, JSAC, and QoSAC, respectively. In Fig. 9(b), the probability that vertical handoff calls will be dropped at the new call arrival rate 0.25 is 0.25, 0.13, and 0.10 in the QoS support, JSAC, and QoSAC, respectively. Both figures show that the QoSAC scheme is more efficient than the other schemes at providing call-level QoS. That it does so is due to the fact that the QoSAC scheme utilizes the resources in each network.

6.3 Revenue performance in the QoSAC

Figure 10 shows the variation of revenue generated against the call arrival rate by the QoSAC scheme in a 24-h period. The comparison is based on the average revenue generated by the QoSAC scheme in each network. When the rate at which new calls arrive increases, a small gap is maintained between the average revenue of the UMTS cellular network and that of the WLAN, due to the load balancing that is achieved by optimal rate allocation. These results show

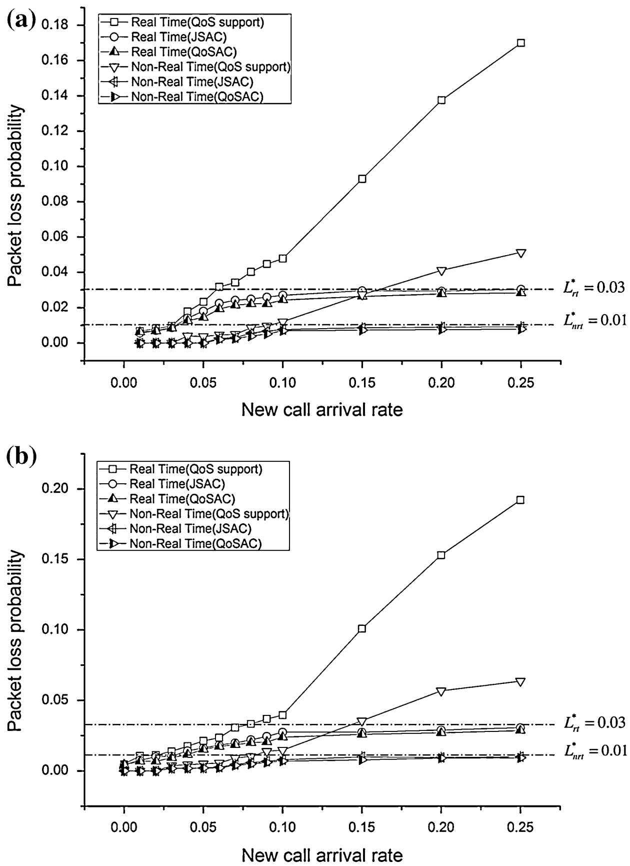

Fig. 8 The probability that a particular packet will be lost in (**a**) the UMTS cellular network (**b**) the WLAN

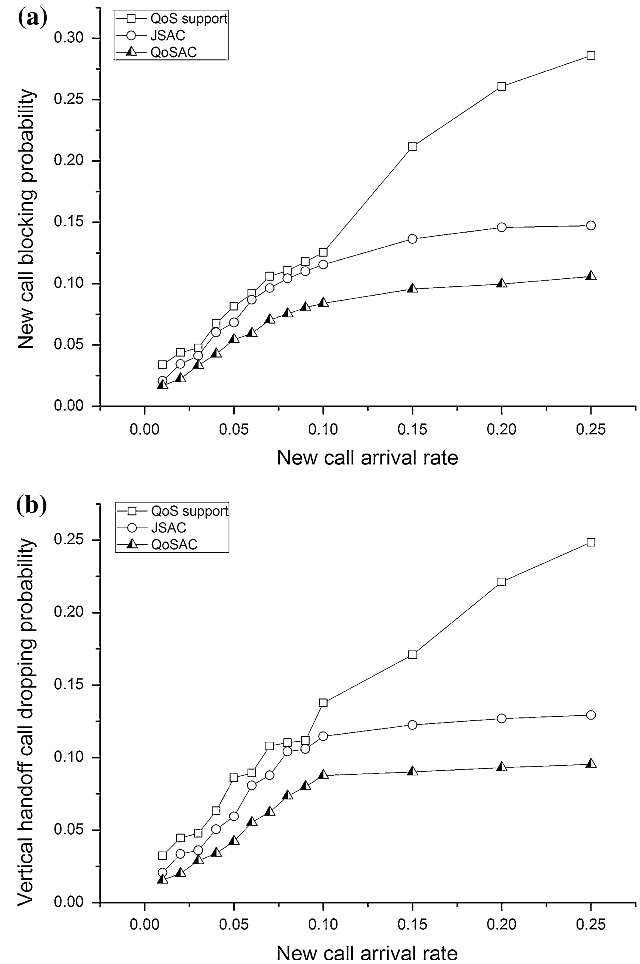

Fig. 9 (**a**) The probability that new calls will be blocked and (**b**) the probability that vertical handoff calls will be dropped

that the QoSAC scheme is very suitable for resource management in a wireless integrated network, due to its strong adaptivity. Sometimes, the WLAN has the greater revenue, whereas at other times, the UMTS cellular network has the greater revenue. This is because more arrival calls access the WLAN before the WLAN is saturated. After the WLAN is saturated, fewer arrival calls access the WLAN because it is not guaranteed to provide QoS. Therefore, the revenue for the WLAN is greater than that for the UMTS cellular network when the call arrival rate is low. However, the revenue for the WLAN is less than that for the UMTS cellular network when the call arrival rate is high.

Figure 11(a) and (b) show the average revenues for QoSAC scheme and the other schemes under various call arrival rates in the UMTS cellular network and WLAN, respectively. In the UMTS cellular network and WLAN, due to the rate allocation technique for the revenue maximization, we can see the significant revenue gain of the QoSAC scheme in high call arrival rates. That is, the average revenues obtained in the QoSAC scheme are 2.2 and 8.0% over than those in the other schemes when the UMTS cellular network and WLAN are in high call arrival rates, respectively. This demonstrates that the CAC with noncooperative rate allocation can achieve better revenue than only vertical handoff provision.

7 Conclusion

We have proposed a new scheme for a wireless integrated network that incorporates network selection into a resource management scheme. The new scheme consists of the mobile stations' preferences, QoS-sensitive admission control, and rate allocation based on game theory. By optimizing the rate allocation to each network, the new scheme is able to balance the load between the networks and maximize network performance under independent and competitive circumstances. The QoSAC scheme also provides the proper unit price for each network and a chance to make admission control QoS-sensitive by making a packet-level metric. Selecting networks according to mobile stations' preferences is more effective and realistic than

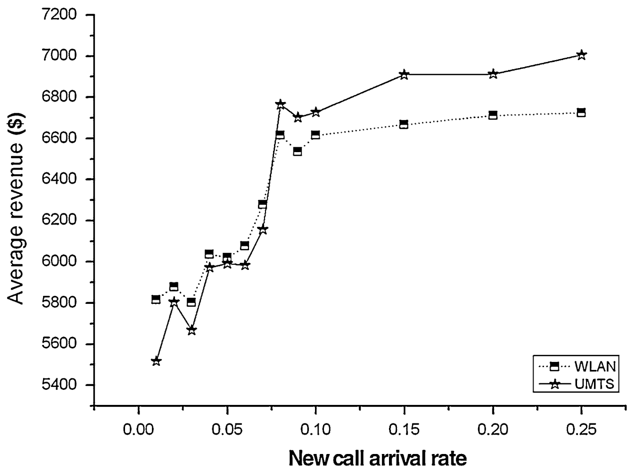

Fig. 10 Revenue comparison between the UMTS cellular network and the WLAN

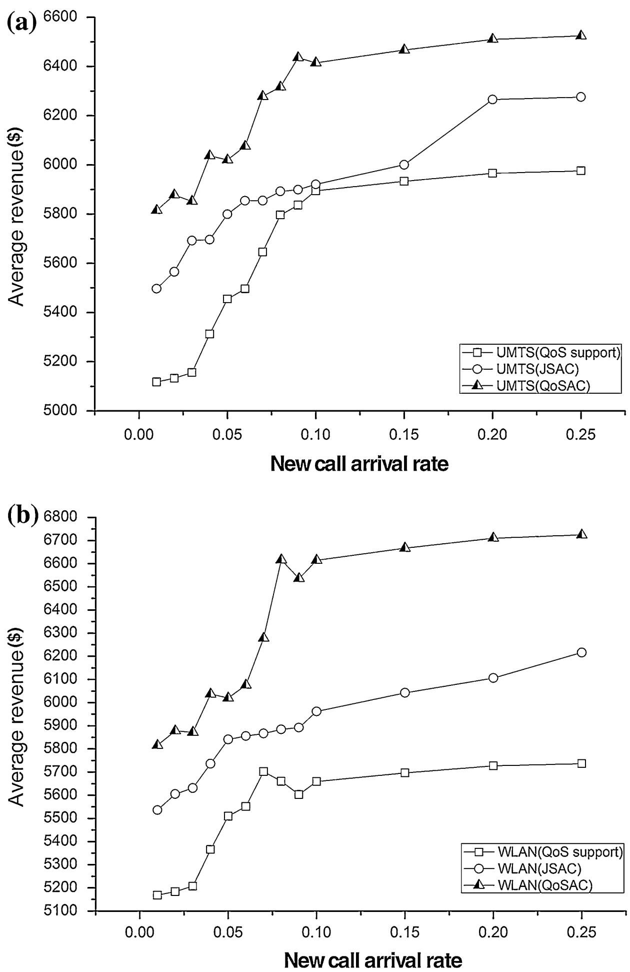

Fig. 11 Revenue comparison between the QoSAC scheme and the other schemes in (**a**) the UMTS cellular network (**b**) the WLAN

selecting according to signal strength. Therefore, the QoSAC scheme provides insights and design guidelines for fourth-generation wireless integrated networks.

In this paper, it was assumed that all mobile terminals are dual-mode terminals. In real environments, the system may also use single-mode mobile terminals. In future work, we will explore the application of the proposed resource management method to environments that have both single-mode and dual-mode mobile terminals. In addition, we will try to obtain an optimal solution for n-person games by adding a newly commercialized network, such as like WiMax. Also, if a network provider operates the two networks, then they can cooperate with each other in order to operate the networks efficiently. Therefore, this cooperative game problem should be studied. It is likely that a better optimal solution can be obtained than the one achieved herein, because mutual negation is possible.

References

1. Mohanty, S. (2006). A new architecture for 3G and WLAN integration and inter-system handover management. *Wireless Networks, 12*(6), 733–745.
2. ETSI, Broadband Radio Access Networks (BRAN); HIPERLAN Type 2; requirements and architectures for interworking between HIPERLAN/2 and 3rd generation cellular systems, TR 101 957.
3. Ruggeri, G., Iera, A., & Polito, S. (2005). 802.11-Based wireless-LAN and UMTS interworking: Requirements, proposed solutions and open issues. *Computer Networks, 47*, 151–166.
4. 3GPP, 3GPP System to WLAN interworking; functional and architectural definition, Technical report 3GPP TR 23.934. v0.3.0, June 2002.
5. Chang, M. F., Wu, L. Y., & Lin, Y. B. (2006). Performance evaluation of a push mechanism for WLAN and mobile network integration. *IEEE Transactions on Vehicular Technology, 55*(1), 380–383.
6. 3GPP TR 22.934. (2003). Feasibility study on 3GPP system to wireless local area network (WLAN) interworking. Release 6, v6.2.0, September 2003.
7. Niyato, D., & Hossain, E. (2007). Integration of WiMAX and WiFi: Optimal pricing for bandwidth sharing. *IEEE Communications Magazine, 45*(5), 140–146.
8. Niyato, D., & Hossain, E. (2008). A noncooperative game-theoretic framework for radio resource management in 4G heterogeneous wireless access networks. *IEEE Transactions on Mobile Computing, 7*(3), 332–345.
9. Song, W., Jiang, H., & Zhuang, W. (2007). Performance analysis of the WLAN-first scheme in cellular/WLAN interworking. *IEEE Transactions on Wireless Communications, 6*(5), 1932–1952.
10. Zhang, Q., Guo, C., Guo, Z., & Zhu, W. (2003). Efficient mobility management for vertical handoff between WWAN and WLAN. *IEEE Communications Magazine, 41*(11), 102–108.
11. Jiang, W. H., Zhuang, W., & Shen, X. (2005). Resource management for QoS support in cellular/WLAN interworking. *IEEE Network, 19*(5), 12–18.
12. Song, W., Zhuang, W., & Cheng, Y. (2007). Load balancing for cellular/WLAN integrated networks. *IEEE Network, 21*(1), 27–33.
13. Yu, F., & Krishnamurthy, V. (2007). Optimal joint session admission control in integrated WLAN and CDMA cellular networks with vertical handoff. *IEEE Transactions on Mobile Computing, 6*(1), 126–139.

14. Suliman, I. M., Oppermann, I., Braysy, T., Konnov, I., & Laitinen, E. (2005). A cooperative multihop radio resource allocation in next generation networks. *IEEE Proceedings of Vehicular Technology Conference, 4*, 2400–2404.
15. Buddhikot, M., et al. (2003). Integration of 802.11, third-generation wireless data networks. *IEEE INFOCOM, 1*, 503–512.
16. Mills, H. (1960). Equilibrium points in finite games. *Journal of the Society for Industrial and Applied Mathematics, 8*(2), 397–402.
17. Epstein, B., & Schwartz, M. (1995). Reservation strategies for multimedia traffic in a wireless environment. *IEEE Proceedings of Vehicular Technology Conference, 1*, 165–169.
18. Goldsmith, A. *Wireless networks*. Cambridge University Press.
19. Wolff, R. W. *Stochastic modeling and the theory of queues*. Prentice-Hall.
20. *IEEE Std. 802.11b*, 1999.
21. Fang, Y., Chlamtac, I., & Lin, Y.-B. (1998). Channel occupancy times and handoff rate for mobile computing and PCS networks. *IEEE Transactions on Computers, 47*(6), 679–692.
22. Lin, Y.-B. (1997). Performance modeling for mobile telephone networks. *IEEE Network Magazine, 11*(6), 63–68.
23. Radunovic, B., & Le Boudec, J.-Y. (2004). Optimal power control, scheduling, and routing in UWB networks. *IEEE Journal on Selected Areas in Communications, 22*(7), 1252–1270.
24. Połowczuk, W. (2003). Pure Nash equilibria in finite two-person non-zero-sum games. *International Journal of Game Theory, 32*(2), 229–240.
25. Saraydar, C. U., Mandayam, N. B., & Goodman, D. J. (2002). Efficient power control via pricing in wireless data networks. *IEEE Transactions on Communications, 50*(2), 291–303.
26. Başar, T., & Olsder, G. T. *Dyanmic noncooperative game theory, 2nd ed*. Society of Industrial and Applied Mathematics.
27. van den Elzen, A. H., & Talman, A. J. J. (1991). A procedure for finding Nash equilibria in bi-matrix games. *Mathematical Methods of Operations Research, 35*(1), 27–43.
28. Binmore, K. *Playing for real: A text on game theory*. Oxford University Press.
29. http://www.nist.gov/.
30. Gang, X., Min, G., Mellor, J. E., Al-Begain, K., & Guan, L. (2005). An adaptive QoS framework for integrated cellular and WLAN networks. *Computer Networks, 47*, 167–183.
31. ITU-T G.114. (1996). One-way transmission time.
32. ITU-T G.1010. (2001). End-user multimedia QoS categories.
33. Holma, H. *WCDMA for UMTS-HSPA evolution and LTE, 4th ed*. Wiley.

Author Biographies

H. M. Shin received the B.S. and M.S. degrees in industrial engineering from Hanyang University, Seoul, Korea, respectively, and Ph.D. degree in industrial & system engineering from Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea, in 2009. He is now a senior researcher at Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. His research interests include modeling, analysis, and optimization in wireless network resource management and network design. He is a member of IEEE Computer Society.

C. Y. Lee is a professor of industrial engineering at Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Korea. He received the B.S. degree in industrial engineering from Seoul National University, Seoul, Korea in 1979, and the M.S. and Ph.D. degrees in industrial engineering from Georgia Institute of Technology, Atlanta in 1981 and 1985, respectively. He is a member of INFORMS and IEEE communication Society. His research area includes mobile communication systems, Internet data communications, heuristic search and optimization. He has published numerous papers in journals related to wireless communications, operations research and optimizations.