Utility-Based Load balancing in WLAN/UMTS Internetworking Systems

Yu Zhou¹, Yanxia Rong¹, Hyeong-Ah Choi¹, Jae-Hoon Kim², JungKyu Sohn³, and Hyeong In Choi³

¹Department of Computer Science, George Washington University, Washington, DC, 20052

²Access Network and Mobile Terminal R&D Center, SK Telecom, Seoul, Korea

³Department of Mathematics, Seoul National University, Seoul, Korea

E-mail:(yuzhou,yxrong,hchoi)@gwu.edu,jayhoon@sktelecom.com,(jgdohn,hichoi)@snu.ac.kr

Abstract—The complementary characteristics of 3G cellular networks (such as UMTS) and IEEE's 802.11 wireless LANs (WLANs) have stimulated intensive research efforts to integrate UMTS and WLAN networks. In such integrated networks, a mobile station is able to switch between UMTS and WLAN networks based on various network parameters such as network throughput, packet delay and other network or user quality criteria. In this paper, we develop a utility-based access selection algorithm targeted to achieve load balancing between UMTS and WLAN networks. The performance of our algorithm is evaluated using OPNET [1] with dual-mode mobile stations.

I. INTRODUCTION

Integration of the IEEE 802.11 wireless LANs (WLANs) and 3G networks, such as Universal Mobile Telecommunication Service (UMTS), has been intensively studied recently due to their complementary characteristics. The 3GPP has been continuously evolving to support multimedia services which require high data rates. Nowadays, a UMTS network can support services with maximum data rate of 2Mbps while a 2.5G cellular network, such as General Packet Radio Service (GPRS), can only provide 100-200 kbps. UMTS networks are gaining popularities and being deployed globally in countries such as UK, Japan and USA. To further increase the data rate at the downlink side, High Speed Downlink Packet Access (HSDPA) was introduced by the 3GPP research community providing data rate up to 14Mbps. Deployment of HSDPA networks has been commercially launched, but the cost for setting up HSDPA infrastructures and replacing existing legacy 3G or 2.5G networks must be a major concern by any service provider. Meanwhile, the commercial success of the IEEE 802.11 protocol makes the access point-based WLAN networks widely deployed in hot-spot areas such as offices, airports and coffee shops. The IEEE 802.11b can provide data rate up to 11Mb/s in 2.4 GHz. The IEEE 802.11a and IEEE 802.11g can provide up to 54 Mb/s in 2.4GHz and 5GHz bands, respectively. But WLANs have disadvantages of small coverage. The coverage by an access point (AP) of WLANs is up to several hundred meters in radius and the cell coverage by

a UMTS Node B is usually several kilometers in radius. Such complimentary characteristics of these two popular networks have stimulated research efforts to integrate UMTS and WLAN networks so that a mobile station can choose a network that has better network quality when it is covered by both networks and receives continuous services over the integrated networks. The hardware requirement for integrating UMTS and WLAN networks is mainly to build dual-mode user equipments (UE) which has capability to access either network. After such a dual-mode UE is available and softwares at each network's operational components are updated, a ubiquitous wireless environment with high data rate enabled in hot spot areas can be set up. The integrated WLAN/UMTS systems can be a good alternative to costly deploying HSDPA systems widely.

In the integrated WLAN/UMTS systems, when a mobile station is covered by both networks, the access control problem arises to decide which network it should be admitted to and when it should switch from one network to the other through vertical handover. One possible architecture is that a decision is made by a new software layer named as IP Switch layer which resides in the UE and keeps monitoring the network situation of current cell. Once the traffic in one network becomes higher and higher and the network efficiency gets impaired, the IP switch layer will deliver the packets from upper layer to the other network interface.

In this paper, we propose a new network access decision algorithm based on the utility-based access control framework. Utility function is a concept borrowed from economics and has been used for scheduling and allocating resources in wireless communication systems, e.g., [12], [13]. In our proposed framework, admission control and vertical handover decisions are made through evaluating utility functions which are provided by the two networks' operational components, i.e., UMTS's Node B, RNC and WLAN's AP. The utility functions are constructed so that each network's capacity is considered to achieve load balancing between UMTS and WLAN networks. Whereas,

the capacity measurements of different networks are various, and it is hard to compare them directly. Therefore, we need a uniform scale to evaluate the network conditions.

The rest of this paper is organized as follows. In the next section, the related work is reviewed. An architecture for dual-mode UEs is described in Section III. In Section IV, our network access decision algorithm is presented. The simulation setup and results are discussed in Section V. Conclusions are given in Section VI.

II. RELATED WORK

Several network architectures for integrated WLAN/UMTS systems have been proposed. The proposed architectures can be grouped into two categories based on the independence between the two networks [14], tight coupling and loose coupling. In the loose coupling architecture, the two networks are integrated beyond the Core Network (CN) of UMTS and are connected through gateways of the Internet. Communications between the two networks are realized through standard IP protocols and the mobility of mobile stations is managed through protocols such as Mobile IP. The loose coupling architecture enables the two networks deployed independently but results in longer delay for signaling and vertical handovers. In the tight coupling architecture, the two networks are integrated at UMTS's CN, which has lower delay for signaling and vertical handover but has higher implementation complexity.

3GPP has been working on standardization for integrating cellular and WLAN systems in [7], [8] in which interworking architecture and interworking scenarios are described.

A policy based access control framework for cellular/WLAN was proposed in [9], where policies are designed to achieve load balancing, but details of the proposed scheme such as performance analysis are not available in [9].

III. SYSTEM MODEL

Inspired by the Dual Mode Terminal (DMT) implemented in [15], we implemented a dual-mode UE (DMUE) in OPNET which can switch between UMTS and WLAN networks. Our DMUE is different from the DMT in which our DMUE can be adopted in loose coupling interworking systems where the UMTS and WLAN networks are connected by a router, whereas DMT is only applicable in tight coupling interworking systems. The protocols in UMTS and WLAN are relatively independent. The packets arriving at the router are routed according to the subnet address of each network. Once the packets are being delivered in the UMTS or WLAN network, the delivery follows individual protocol. The main difference of UMTS and WLAN mobile station is in the MAC layer and the physical layer. In the DMUE, we created a new software layer, called IP switch layer, just below the IP layer. Each DMUE has multiple IP addresses which access to different

networks. When a data packet comes from the IP layer, the IP switch layer will make a network access decision based on the utility functions and send out this data packet through corresponding physical layers.

IV. NETWORK ACCESS DECISION ALGORITHMS

The stations are admitted/handovered to the network that has higher utility, i.e., available bandwidth. Thus, the resources of the integrated networks are fully utilized, and meanwhile the load balancing among networks is achieved. Next we discuss how utility functions are designed for UMTS and WLAN, respectively.

A. Utility function for the UMTS

In the UMTS networks, the admission control procedure is started when a new service is requested. The request includes traffic's QoS requirement such as data rate, delay requirement, etc. After the UTRAN (RNC and Node B) of the UMTS network receives the request, it will decide whether to grant the request based on the network condition. The network condition is evaluated by the UTRAN through computing *load factors* for uplink and downlink [16]. The load factor for the uplink is defined as:

$$\eta_{UL} = (1 + i) \cdot \sum_{j=1}^N \frac{1}{1 + \frac{W}{(E_b/N_o)_j \cdot R_j \cdot v_j}} \quad (1)$$

where N is the number of stations, v_j is the active factor of station j at physical layer, E_b/N_o is the signal energy per bit divided by noise spectral density, W is the chip rate, R_j is the bit rate of station j and i is the other cell to own cell interference ratio seen by the base station receiver. The load factor for the downlink is

$$\eta_{DL} = \sum_{j=1}^N v_j \cdot \frac{(E_b/N_o)_j}{W/R_j} \cdot [(1 - \alpha_j) + i_j], \quad (2)$$

where α_j is the orthogonality of channel of station j and i_j is the ratio of other cell to own cell base station received by user j .

In UMTS systems, load factors are always controlled to be below than a threshold, say η_{max} ($\eta_{max} < 1$). Most systems have specific values for η_{max} .¹ When a service request comes, the UTRAN estimates the new resulting load factors for both uplink and downlink. The UTRAN will accept the request only when both the new resulting downlink and uplink load factors are smaller than η_{max} .

Given the current uplink load factor, η_{UL} , a simple utility function $f_{UL}(\eta_{UL})$ can be constructed as

$$f_{UL}(\eta_{UL}) = B_{UL} \cdot (1 - \frac{\eta_{UL}}{\eta_{max}}), \quad (3)$$

¹In most UMTS systems, 0.75 is a value commonly used for both uplink and downlink. In this paper, we assume the uplink and downlink load factors have the same value of η_{max} .

where B_{UL} is the maximum uplink bandwidth (in bits per second).

Correspondingly, given the current downlink load factor, the linear utility function for downlink, $f_{DL}(\eta_{DL})$, is:

$$f_{DL}(\eta_{DL}) = B_{DL} \cdot (1 - \frac{\eta_{DL}}{\eta_{max}}). \quad (4)$$

where B_{DL} is the maximum downlink bandwidth (in bits per second).

Thus, the utility function of the uplink/downlink provides an estimation how much bandwidth is available in the uplink/downlink. Then the utility function for the UMTS network is designed as

$$f_{UMTS} = f_{UL}(\eta_{UL}) + f_{DL}(\eta_{DL}). \quad (5)$$

B. Utility function for the WLAN

A WLAN network has much more bandwidth than a UMTS network does. A desirable scenario in the integrated WLAN/UMTS networks would be that in a hotspot area, i.e., covered by WLAN, most of the stations are connected to the WLAN to enjoy the high data rate of WLAN, while in the area outside hotspots, i.e., only covered by UMTS, static or mobile stations are connected to UMTS to enjoy the large coverage of UMTS.

However, WLAN does not have explicit QoS control. When the WLAN is heavily loaded, some QoS metric such as delay cannot be guaranteed. Moreover, as pointed out in [11], WLAN achieves less throughput when the network is saturated than that when the network is not saturated. As the traffic load (number of stations) increases, severe collisions occur, which results in that the stations can barely transmit a packet successfully. Thus, the WLAN network should be closely monitored such that the network is not *overbusy*. An indicator reflecting the WLAN utilization adopted in the literature is *busyness ratio* [11], which is defined as the ratio of the time that the network is sensed busy. So stations are admitted/handovered to a WLAN network only when its busyness ratio R_b is less than a threshold, say R_{th} . Given the current busyness ratio R_b and its upper bound R_{th} , the utility function for WLAN indicating the available bandwidth to accommodate new stations is given as

$$f_{WLAN} = \frac{EP(R_{th} - R_b)}{T_s}, \quad (6)$$

which is simply derived from [11]. In equation (6) EP is the average DATA packet size and T_s is the average time associated with a successful transmission. For the case that RTS/CTS is not used,

$$T_s = DIFS + T[EP] + SIFS + ACK, \quad (7)$$

and for the case that the RTS/CTS mechanism is used,

$$T_s = DIFS + RTS + SIFS + CTS + T[EP] + SIFS + ACK \quad (8)$$

where $T[EP]$ is the time for transmitting a packet with payload size EP , and RTS , CTS and ACK are the time for transmitting the RTS, CTS and ACK packet, respectively. Note that f_{UMTS} and f_{WLAN} both denote remain bandwidth in bps.

C. Network access decision

Each AP of WLAN and Node-B of UMTS calculate their own utility functions (either periodically or triggered by events). The computed utilities are then broadcasted. Once a station receives the utility from either UMTS or WLAN network, it will compare the received utility with the utility of the host network to decide whether to switch network.

Algorithm NetworkAccessDecision()

```

1: if the utility comes from UMTS then
2:   new_utility =  $f_{UMTS}$ ;
3: else
4:   new_utility =  $f_{WLAN}$ ;
5: end if
6: new_network_id = the id of the network where the new
   utility comes from;
7: if current_network_id = new_network_id then
8:   current_utility = new_utility;
9: else
10:  if new_utility > current_utility + utility_gap then
11:    switch to the network with id being new_network_id;
12:  end if
13: end if

```

Notice that to avoid unnecessary *oscillation*, i.e., a station keeps switching back and forth between two networks, a variable, *utility_gap*, is introduced such that only when the utility of a candidate network is larger than the utility of host network by *utility_gap*, the station changes the network. Secondly, as the network utility is broadcasted, all stations will receive it at the same time (if the transmission time in the media is ignorable). Then, all the stations will try to switch to the network that has higher utility. As a result, the network that has higher utility before will be loaded very quickly (i.e., low utility) and the network that has lower utility before will be depleted (i.e., high utility), which lead the stations to switch the network again. To avoid this undesirable *network trembling*, each station keeps a random number T_{stay} . Each station has to stay in a network for at least T_{stay} seconds before switching to another network.

V. SIMULATION RESULTS

In the simulation scenario, we create a UMTS network with one Node B and a WLAN with one AP in OPNET as shown in Figure 1, and 10 DMUEs in the intersection coverage area of the two networks. Each DMUE has two IP addresses, one is for UMTS, the other is for WLAN. An application of UDP video uploading from each DMUE to the server is applied. Each video frame size is 34560 bytes, and it has an interarrival time with a poisson distribution. The total simulation time is 1200 seconds.

Fig. 1. System Model

In case 1, all the DMUEs initially connect to WLAN network, the network access decision algorithm based on the utility functions we proposed works in the whole simulation duration.

In case 2, in the initiation, all the DMUEs randomly connect to either UMTS network or WLAN network, and switchings never happen.

The parameters for utility functions and other parameters are summarized in Table 1:

TABLE I
SUMMARY OF PARAMETERS

Parameters	Values
Simulation Time	1200 seconds
B_{UL}, B_{DL}	0.5 Mbps
WLAN Data Rate	1 Mbps
$SIFS$	10 μs
$DIFS$	50 μs
Phy header	192 bits
MAC header	224 bits
DATA packet	12000 bits + Phy header + MAC header
RTS	160 bits + Phy header
CTS, ACK	112 bits + Phy header
R_{th}	0.8
μ_{max}	0.75
Traffic frame size	34560 bytes
Traffic frame interarrival time	poisson(2) seconds
utility_gap	10000 bits

From the simulation result shown in Figure 2 which is measured at the server side, we see that the overall throughput with utility based approach is larger than the one with random access. In the former case, the throughput has a sharp drop in the beginning due to the abrupt increment of traffic in the network, which causes collisions when all DMUEs trying to get admission by the other network. But when the overall network throughput becomes stable, the drawback will be counteracted by our utility-based network access algorithms.

VI. CONCLUSION

In this paper, we propose a utility based access control framework for an integrated WLAN/UMTS dual system. In this framework, the UMTS's UTRAN and the WLAN's AP measure their respective network conditions through available bandwidth and reflect network conditions numerically by calculated utilities dynamically. Each network's utility

Fig. 2. Throughput in the system

is broadcasted to every DMUE. The DMUE's handover decisions are made by comparing the received utilities and switch to the network with better utility. The preliminary results show that our proposed scheme performs significantly better than the random selection. Further analysis remains to be done.

REFERENCES

- [1] OPNET network simulator, <http://www.opnet.com/>.
- [2] IEEE 802.11 WG, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specification, Standard, IEEE, Aug. 1999.
- [3] IEEE 802.11b WG, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specification: High-speed Physical Layer Extension in the 2.4 GHz Band, IEEE, Sep. 1999.
- [4] IEEE 802.11a WG, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specification: High-speed Physical Layer Extension in the 5 GHz Band, IEEE, Sep. 1999.
- [5] IEEE 802.11g, Further Higher-Speed Physical Layer Extension in the 2.4 GHz Band, 2003.
- [6] 3GPP, 3rd Generation partnership Project; Technical Specification Group Radio Access Network; RRC Protocol Specification for Release 1999. Technical Specification 3G TS 25.331 version 3.5.0 (2000-12), 2000.
- [7] 3GPP TK 22.934, Feasibility study on 3GPP system to wireless local area network (WLAN) interworking, v. 1.0.0, Release 6, Feb. 2002.
- [8] 3GPP, Group Services and System Aspects; 3GPP Systems to Wireless Local Area Network (WLAN) Interworking; System Description (Release 6), TS 23.234, v. 6.2.0, Sept. 2004.
- [9] W. Song, W. Zhuang, and Y. Cheng, "Load Balancing for Cellular/WLAN Integrated Networks," *IEEE Network*, issue 1, 2007.
- [10] Giuseppe Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," *IEEE Journal on Selected Areas in Commun.*, vol. 18, no.3, March 2000.
- [11] H. Zhai, X. Chen, and Y. Fang, "How Well Can the IEEE 802.11 Wireless LAN Support Quality of Service", *IEEE Trans. Wireless Commun.*, vol. 4, no. 6, Nov. 2005.
- [12] R. Agrawal, V. Subramanian and R. Berry, "Joint Scheduling and Resource Allocation in CDMA Systems", *Proc. of 2nd Workshop on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt 04)*, Cambridge, UK, March 24-26, 2004.
- [13] J. Huang, V. Subramanian, R. Agrawal, and R. Berry, "Downlink Scheduling and Resource Allocation for OFDM Systems," *Conf. on Information Sciences and Systems (CISS)*, March 2006.
- [14] M. Buddhikot, G. Chandranmenon, S. Han, Y. W. Lee, S. Miller and L. Salgarelli, "Integration of 802.11 and Third-Generation Wireless Data Networks," *Proc. IEEE INFOCOM*, 2003.
- [15] F. Siddiqui, S. Zeadally and S. Fowler, "A Novel Architecture for Roaming between 3G and Wireless LANs", *1st Int'l Conf. on Multimedia Services Access Networks, MSAN'05*, 2005.
- [16] H. Holma, A. Toskala, "WCDMA for UMTS Radio Access for Third Generation Mobile Communications," *John Wiley & Sons, 3rd edition*, 2004.