

Ubiquitous Wireless Connectivity across Cellular and Wireless Local Area Networks*

F. SIDDIQUI AND S. ZEADALLY[†]

*Department of Computer Science
Wayne State University
Detroit, MI 48201, U.S.A.*

[†]*Department of Computer Science and Information Technology
University of the District of Columbia
Washington, DC 20008, U.S.A.*

Future generation wireless networks are expected to support multiple wireless access technologies, each with different access bandwidth and coverage range. Two of these technologies include Universal Mobile Telecommunications System (UMTS) and IEEE 802.11 Wireless Local Area Networks (WLANs). Provision of real time ubiquitous communication to mobile users, anywhere, anytime through these heterogeneous networks remains a challenge. This paper proposes an end-to-end mobility management solution that enables mobile users to maintain seamless connectivity while moving across such heterogeneous access networks. The continuous network connectivity is obtained through the use of mobile devices equipped with dual network interfaces and the capability to switch data transmission between these interfaces depending upon the availability of the access network. We present and discuss the design of such a dual-mode radio access device and propose a simple inter-technology handoff technique. We evaluate the performance of our proposed architecture by conducting experimental simulation tests using OPNET.

Keywords: wireless, ubiquitous, UMTS, handoff, cellular

1. INTRODUCTION

Mobile communications and wireless networks are developing at a rapid pace. Advanced techniques are emerging in both these disciplines. There exists a strong need for integrating WLANs with 3G networks to develop hybrid mobile data networks capable of data services and very high data rates in hotspots. 3G wireless systems such as Universal Mobile Telecommunication Systems (UMTS) can provide mobility over a large coverage area, but with relatively low speeds of about 144 Kbits/sec. On the other hand, WLANs provide high speed data services (up to 11 Mbits/sec with 802.11b) over a geographically smaller area. WLANs are generally used to supplement the available bandwidth and capacity of a 3G network in hotspot areas such as railways and airports with high traffic-densities, without sacrificing the capacity provided to cellular users.

The rest of this paper is organized as follows. Section 2 provides a brief background on 3G and WLAN networks. Section 3 presents a comparison of the various internet-working architectures. In section 4 we compare two integration architectures connecting

Received December 16, 2005; revised March 2 & May 4, 2006; accepted May 11, 2006.

Communicated by Ten-Hwang Lai.

* This work was supported by grants from Microsoft (Seattle), OPNET (Maryland), and Ixia Corporation (Calabasas).

UMTS and 802.11 networks. Section 5 discusses the design and simulation of the proposed switching and handoff techniques for roaming between 3G and WLAN networks. In section 6 we compare our work with previous approaches aimed at evaluating and developing 3G-WLAN integration schemes. Finally, in section 7, we present some concluding remarks.

2. BACKGROUND

2.1 WLAN Standards

802.11b [1] WLAN has been widely deployed in offices, homes and public hotspots such as airports and hotels given its low cost, reasonable bandwidth (11Mbps/s), and ease of deployment. However, a serious disadvantage of 802.11 is the small coverage area (up to 300 meters) [2]. Other 802.11 standards include 802.11a and 802.11g, which allow bit rates of up to 54 Mbits/sec.

2.2 3G Networks

ITU defines 3G as any device that can transmit or receive data at 144 Kbps or better [3]. In practice, 3G devices can transfer data at up to 384 Kbps. As a comparison, Global System for Mobile Communications (GSM) data rates are up to 14.4 Kbps and General Packet Radio Service (GPRS) is around 53.6 Kbps used in 2G and 2.5G respectively. Two main proposed systems for 3G recognized by the International Telecommunication Union (ITU) are *Code Division Multiple Access (CDMA 2000)* and *Universal Mobile Telecommunication System (UMTS)* [4]. UMTS is composed of two different but related modes: Wideband CDMA also called Frequency Division Duplex (FDD) and CDMA-Time Division Duplex (TDD). FDD mode is considered the main technology for UMTS. Separate 5 MHz carrier frequencies are used for the uplink and downlink respectively, allowing an end-user data rate up to 384 Kbits/sec. The TDD mode is Time-Division Synchronous Code-Division Multiple-Access (TD-SCDMA). TD-SCDMA operates on low-chip-rate carriers and allows end-user data rates up to 2Mbits/sec.

2.3 3G Network Architecture

A 3G network (Fig. 1) consists of three interacting domains- a Core Network (CN), Radio Access Network (RAN) and the User Equipment (UE) also called a 3G mobile station [3, 13]. 3G operation utilizes two standard suites: UMTS and CDMA2000 which have minor differences with respect to the components they have in the RAN and the CN. The main function of the 3G core network is to provide switching, routing and transit for user traffic. It also contains the databases and the network management functions. The core network is divided into Circuit-switched (CS) and Packet-switched (PS) domains. Circuit switched elements include Mobile Services Switching Center (MSC), Visitor Location Register (VLR), and gateway MSC. These circuit switched entities are common to both the UMTS as well as the CDMA2000 standards.

The differences in the CN with respect to the two standards lie in the PS domain. Packet-switched elements in UMTS include Serving GPRS Support Node (SGSN) and

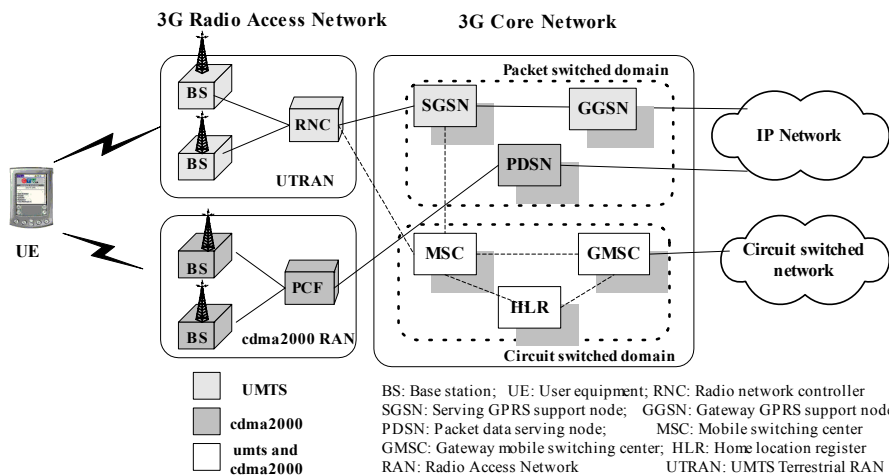


Fig. 1. Components of a 3G network.

Gateway GPRS Support Node (GGSN). CDMA2000 packet-switched component is primarily the Packet Data Serving Node (PDSN). Some network elements like Equipment Identity Register (EIR), Home Location Register (HLR) are shared by both domains. The main function of the MSC server is to handle call-control for circuit-based services including bearer services, *etc.* The MSC server also provides mobility management, connection management and capabilities for mobile multimedia as well as generation of charging information. It can also be co-located with the Visitor Location Register (VLR).

GGSN is the gateway to external data networks. It supports control signaling towards external IP networks for authentication and IP-address allocation, and mobility within the mobile network. GGSN provides functions for forwarding and handling user information (IP packets) to and from external networks (Internet/intranets). SGSN provides session management, *i.e.* mechanisms for establishment, maintenance and release of end user Packet Data Protocol (PDP) contexts. It also provides mobility management and supports inter-system handoff between mobile networks. SGSN also supports generation of charging information. The PDSN incorporates numerous functions within one node. Routing packets to the IP network, assignment of dynamic IP-addresses and maintaining point-to-point protocol (PPP) sessions are some of its main functions. It also initiates the authentication, authorization and accounting (AAA) for the mobile station. The radio access network provides the air interface access method for the user equipment. An UMTS RAN (UTRAN) consists of Radio Network Controllers (RNC) and Base Stations (BS) or Node-B. The RNCs manage several concurrent Radio Link Protocol (RLP) sessions with the User Equipments and per-link bandwidth management. It administers the Node-B for congestion control and loading. It also executes admission control and channel code allocation for new radio links to be established by the Node-B.

A CDMA2000 RAN consists of a base station and 2 logical components- the Packet Control Function (PCF) and the Radio Resources Control (RRC). The primary function of the PCF is to establish, maintain and terminate connections to the PDSN. The PCF communicates with the RRC to request and manage radio resources in order to relay packets to and from the mobile station. It also collects accounting information and for-

wards it to the PDSN. RRC supports authentication and authorization of the mobile station for radio access and supports air interface encryption to the mobile station.

3. WLAN AND 3G CELLULAR DATA NETWORK INTEGRATION ARCHITECTURES

3G and WLAN Networks have been integrated with different approaches and strategies. The two most commonly used approaches are *tight* and *loose* internetworking. Other strategies are modifications of these two basic approaches.

In the tight internetworking approach, the WLAN network does not appear to the 3G core network as an external packet data network. Instead, it appears simply as another 3G Radio Access Network. WLAN network in this case emulates several functions of the 3G RAN. This is made possible by employing a specialized WLAN gateway in between the 3G core network and the WLAN network that hides the details of the 802.11 network and implements all 3G protocols required in the 3G Radio Access Network. The architecture is capable of providing roaming services to users with dual stack (3G and 802.11) network cards in their mobile devices. Using this approach, both the networks often share common billing and authentication mechanisms. However, all traffic from the WLAN network passes through the 3G core network, which could cause it to become a bottleneck. Tight internetworking also requires common ownership of the two networks that does not make it a very feasible deployment strategy. However, tight internetworking does offer high security mechanisms as the 3G security protocols can be reused in the WLAN network. It also provides fast handoffs as roaming between the two networks is the same as moving between two RANs of the same 3G network (since the WLAN network appears as a different Routing Area only).

Loose internetworking, on the other hand has the WLAN gateway directly connected to the Internet and does not have any direct link to 3G network elements. In contrast to tight coupling, the WLAN data traffic does not pass through the 3G core network but goes directly to the IP network (Internet). In this approach, different mechanisms and protocols can handle authentication, billing and mobility management in the 3G and WLAN portions of the network. Loose coupling has low investment costs and permits independent deployment and traffic engineering of the WLAN and 3G networks. Widespread coverage is possible through roaming agreements with many partners. However, roaming in this setup suffers from high handoff delays and service disruption periods.

Other 3G-WLAN internetworking strategies and their features are summarized in Table 1. The Mobile IP [5] internetworking architecture considers 3G and WLANs as independent networks. It allows easy deployment but suffers from long handoff latency and might not be able to support real-time services and applications. The gateway approach [10] permits independent operation of the two networks and provides seamless roaming facility between them. It connects the two networks using a new logical node called the virtual GPRS support node. The gateway approach does not require the use of Mobile IP and has a comparatively lower packet loss during handoff. The emulator approach [5] is difficult to deploy since it requires combined ownership of the two networks but does yield low handoff latency and is thus well suited for real-time applications. Finally, the peer-networks [12] approach allows easy deployment but also suffers

Table 1. Comparison of 3G-WLAN internetworking strategies.

Internetworking Approach	Handoff Latency	Network Management and Ownership	Authentication	Billing	Real time application support
Mobile-IP approach	High	Separate ownership of 3G and WLAN networks permitted with roaming agreements	Reuse 3G cellular AAA functions in WLANs	Uses 3G billing feature	No
Gateway approach	Low	Permits 3G and WLAN networks to operate independently	Can provide both, separate WLAN security as well as 3G AAA	Can provide separate as well as combined billing mechanism	Yes
Emulator approach	Low	Combined ownership required	Uses 3G authentication mechanism	Uses 3G billing mechanism	Yes
Tight coupling	Low	Generally requires both WLAN and 3G networks to be owned by same operator	3G ciphering key used for WLAN encryption	3G accounting features re-used	Yes
Loose coupling	High	Permits independent deployment of WLAN and 3G networks.	Cellular access gateway provides authentication	Billing mediator to provide common accounting	Not very suitable
Peer networks approach	High	Both, same or different operator ownership permitted	Use of AAA functionality of the 3G UMTS network	Billing feature of 3G UMTS network used	Not very suitable

from high handoff delays, thereby making it unsuitable for real-time applications. The choice of the integration architecture is important since multiple integration points exist with different cost- performance benefits for different scenarios.

4. INTEGRATING UMTS AND IEEE 802.11 NETWORKS: SGSN VS. GGSN

We evaluated (via simulations using OPNET [11]) two internetworking architectures to interoperate the 3G (UMTS) and WLAN networks by connecting them at two strategic points – the **SGSN node** and the **GGSN node** as shown in Fig. 2.

4.1 Integrating UMTS and IEEE 802.11 Networks

We evaluated via simulations using OPNET [10], two internetworking architectures that enable interoperability between 3G (UMTS) and WLAN networks by connecting them at two strategic points namely, the SGSN node and the GGSN node (as shown in Fig. 2). In a real scenario, a single operator's network may consist of several WLANs. Depending upon how the WLANs are connected with the UMTS core network, the network operator may deploy various numbers of SGSNs and GGSNs to achieve scalability and load balancing. In this work, we are interested in comparing the performance of two WLAN-UMTS integration methods (using SGSN and GGSN as the integration points) for determining the latency experienced during communication between WLAN and UMTS users. In the GGSN integration mode, the WLAN access point can be connected to the GGSN via a direct link or through an intermediate ISP's network over the public

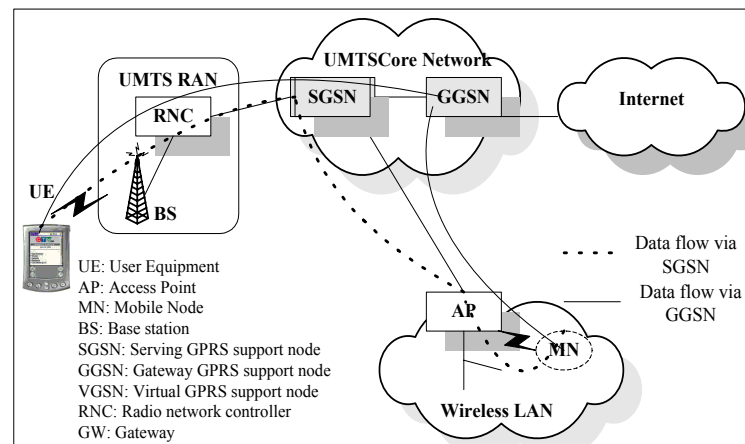


Fig. 2. UMTS-WLAN integration at SGSN and GGSN.

Internet. However, in this simulation we have assumed the connections for both SGSN and GGSN integration to be direct links in order to compare the performance of the integration points (with respect to the different communication techniques used in each case, such as the SGSN integration uses a specialized WLAN access point and GMM attach procedures, while the GGSN integration uses the GPRS tunneling protocol) and eliminate any other factors such as extra hop counts, *etc.* in any scenario which may affect the latency results. Thus for the purpose of using a comparable network configuration, we assume both SGSN as well as GGSN to have a direct link with the WLAN access points.

When the UMTS and WLAN networks are connected through the SGSN node, the WLAN network does not appear to the UMTS core network as an external packet data network. Instead, it simply appears as another radio access network. The WLAN AP in this case needs to have the capability of processing UMTS messages. Thus, whenever a Mobile Node (MN) in the WLAN network wants to exchange data with the UMTS User Equipment (UE), it first needs to undergo the GMM attach procedure to notify the SGSN of the location on the communicating node and also to establish a packet-switched signaling connection. The WLAN AP is responsible for sending these request messages to the SGSN on behalf of the WLAN MN.

The GMM attach procedure is a three-way handshake between the MN, RNC and the SGSN. Upon completion of this procedure, the WLAN MN is authenticated into the UMTS network.

In the GGSN integration case, whenever a MN in a WLAN network wants to communicate with a UE in the UMTS network, it does so through the GGSN node. The UE in the UMTS network first activates the Packet Data Protocol (PDP) context that it wants to use. This operation makes the UE known to its GGSN and to external data networks, in this case, the WLAN network. User data is transferred transparently between the UE and the WLAN network with a method known as encapsulation and tunneling. The protocol that takes care of the encapsulation is the GPRS Tunneling Protocol (GTP). For this kind of internetworking configuration, the WLAN AP is a simple 802.11b access point and does not need to process UMTS messages.

4.2 Evaluation of UMTS-WLAN Integration Architectures: SGSN vs. GGSN

Our simulation setup consisted of a UMTS network connected to a WLAN network using two different integration points (SGSN and GGSN) as shown in Fig. 2. Parameters such as end-to-end voice packet delay, file upload times, and web-page retrieval times were measured to evaluate these two integration architectures with respect to communication delays involved when users exchange data across these networks. This is in contrast to the previous research efforts which focused only on measurement of handoff delays experienced when users moved from one network to another while accessing a resource from the external IP data network.

We conducted various tests to measure the various application delays involved when users directly communicate across the UMTS-WLAN networks without actually accessing any external data networks. The goal of this performance evaluation is to compare the delays involved when user data is exchanged across the UMTS and WLAN networks connected at two different points, namely, the GGSN and the SGSN. Various types of applications were used to evaluate the performance of two different integration architectures as well as our dual-mode radio access protocol design. The application types used included: Voice over IP (VoIP) in GSM encoded format, FTP, and HTTP-Web browsing (as shown in Table 2). The parameters measured include end-to-end packet delay, file upload time, HTTP page response time, and handoff delay when moving from one network to another. Other properties associated with each application are summarized in Table 2.

Table 2. Simulated applications.

Application	QoS Class	Measurement Parameter	Size	Protocol
FTP	Background	File download time	100-1000 Kilobytes	TCP
FTP	Background	File upload time	100-1000 Kilobytes	TCP
GSM encoded voice	Conversational	End-to-end delay	33 Bytes	UDP
GSM encoded voice	Conversational	Jitter	33 Bytes	UDP
HTTP Web browsing	Interactive	Page response time	3000 Bytes	TCP

Simulations performed for both UDP and TCP flows are presented. For the UDP flow (VoIP traffic), end-to-end packet delays and jitter were measured. For TCP flows (FTP, HTTP) the upload/download response times were measured. Fig. 3 shows the average file download and upload times experienced when transferring files of various sizes between the User Equipment and the WLAN Mobile Node under two different integration scenarios. We can see that the download/upload times are much lower for all file sizes when the transfer is done through the GGSN node as compared to when the networks are interconnected at the SGSN. Table 3 shows the average delay and jitter for VoIP packets and the response time in the case of HTTP. It is observed that both, delay and jitter values are much lower in the GGSN case. The average end-to-end delay (using a packet size of 33 bytes) is approximately 0.05 second (50 ms) in the GGSN case, which is within the range of tolerable delay (0 to 150 ms) [11] while delay for the SGSN case is much higher (as shown in Table 3).

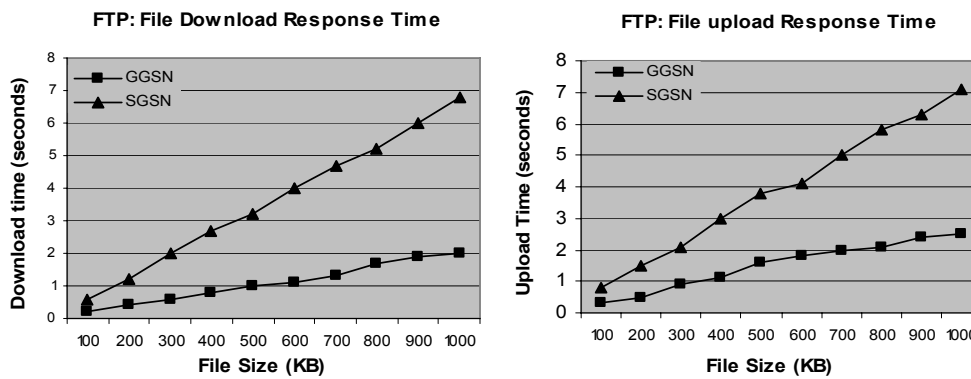


Fig. 3. FTP file download/upload time.

Table 3. Application response times (seconds) for VoIP and HTTP (seconds).

Integration Point	VoIP delay	VoIP jitter	HTTP page response time
SGSN	0.45	0.15	14.7
GGSN	0.05	0.03	4.8

The simulation results reveal that the application response time (delay) is consistently higher in the case where the UMTS and WLAN networks are connected through the SGSN node, as compared to the case where the two networks are connected at the GGSN. This higher response time can be attributed to the additional processing time required at the WLAN access point in the first case. When the two networks are connected at the SGSN node, the WLAN access point has to perform the functions of a RNC on one interface and on the other interface implements the functionality of a WLAN AP. Therefore, it has to perform the additional initialization steps to authenticate the WLAN Mobile Node to the UMTS network, in the form of GMM Attach procedure and PDP context activations. Furthermore, internetworking through the SGSN requires the WLAN access point to be UMTS-aware and process each packet from the WLAN Mobile Node by adding the required Packet Data Protocol (PDP) headers before it can be sent to the SGSN. This adds additional overhead to each packet. When integration is done at the GGSN node, the WLAN AP is a simple 802.11b access point and does not require any special capabilities to process UMTS messages. Data packets are transferred between the User Equipment and the WLAN network using encapsulation by the GPRS tunneling protocol. This reduces the packet latency, as there is no additional delay due UMTS initialization procedures or packet conversions.

It is worthwhile mentioning that, in the simulations presented, the GGSN is one hop away from the SGSN. We have achieved lower communication delays in the GGSN integration case because the additional delay added due to the extra hop in the case of the GGSN is comparatively lower compared to the delay introduced due to the additional processing required at the specialized WLAN access point used to connect to the SGSN. More specifically, when the WLAN and UMTS networks are connected through the GGSN, the GGSN connects to the WLAN access point using the "Gi" interface. This makes the GGSN appear to the WLAN as a simple router. Therefore, whenever a mobile node in the WLAN wants to communicate with a node in the UMTS network, it simply

routes packets through the GGSN. It does not need to authenticate itself to the UMTS network or undergo any GMM attach procedures.

In the SGSN integration case, the SGSN connects to the WLAN via the “Gb” interface. This makes the WLAN appear to the SGSN as another Routing Area (RA) of the 3G (UMTS) network. Whenever a mobile node in the WLAN wants to communicate with a node in the UMTS network, it needs to authenticate itself to the UMTS core network and undergo GMM Attach procedures and PDP context activations. These signaling procedures introduce additional delays when using SGSN integration for communication between mobile users in WLAN and UMTS networks. Thus, the communication latency observed was lower when users in the WLAN and UMTS networks communicated through the GGSN.

The advantages, however, of using SGSN integration scheme include the reuse of UMTS authentication, authorization, accounting (AAA) mechanisms, usage of common subscriber databases and billing systems, increased security features (since the UMTS security mechanisms are reused), as well as possibility of having continuous sessions as users move across the two networks, since handoff in this case is very similar to an intra-UMTS handoff as the WLAN AP appears as another RNC to the SGSN node. In the case of GGSN integration, since the WLAN is considered to be an external network, different billing and security mechanisms are needed.

In the following section, we discuss the design and implementation of a novel UMTS-WLAN integration scheme based upon the GGSN integration method, which allows low application response times as well as seamless mobility across the two networks.

5. DESIGN OF DUAL-MODE TERMINAL AND HANDOFF MECHANISM

A network simulation model was constructed using OPNET 10.0.A simulator [10]. Fig. 4 illustrates the network level view of the simulated network topology. It is composed of a WLAN network (hotspot) of radius 150 meters located within a UMTS cell. The two networks are integrated at the GGSN node. The topology chosen for our simulation tests corresponds to real-world scenarios where a WLAN network serves as a hotspot within a 3G network. Such hotspots can be found in office buildings, hotels, airports, train/bus stations, cafés *etc.* WLAN is a low cost broadband technology; however it does not provide end-to-end coverage. The topology can be a good example of a scenario where a corporate user uses the WLAN in his office building and as he steps out of the building, he uses the 3G network (in this case UMTS). The WLAN network is within the coverage of a 3G network. Another example that reflects such a setup is as follows: a typical user on the way to his office might roam from a home WLAN network onto a 3G network then onto a public WLAN hotspot, say in a café, and then back onto a 3G network and then finally onto an enterprise WLAN. Thus the description of the 3G-WLAN architecture simulated topology closely represents actual user mobility scenarios. The different entities of the two networks (WLAN and 3G) described in the topology are however, generic to all cases. The UMTS network is composed of a UTRAN (RNC and Node B) and a UMTS core network (SGSN and GGSN). Details of 3G network components are outlined in [3].

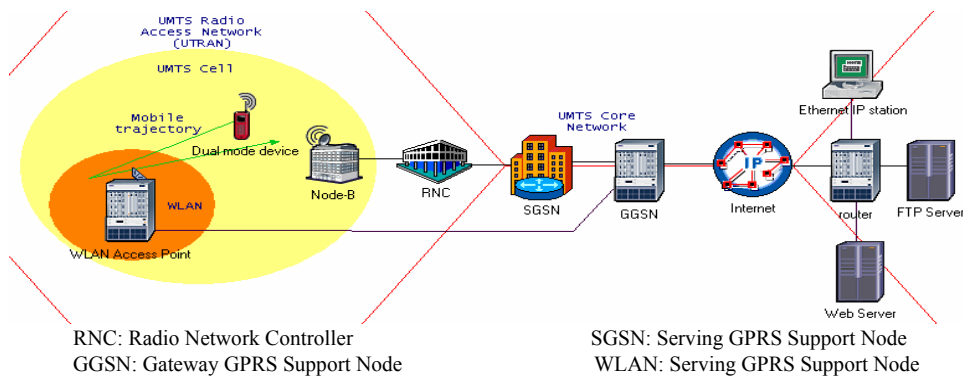


Fig. 4. 3G-WLAN network topology.

In the simulated network topology, the mobile node is a Dual Mode Terminal (DMT) with both a UMTS interface as well as a WLAN interface. We have designed and implemented a software layer called the **Switching Module** incorporated into the protocol stacks. The Switching Module selectively switches the access interface between the UMTS radio stack and the WLAN radio stack depending on the network availability. The Switching Module also makes intelligent interface selection in the presence of overlapped coverage between UMTS and WLAN networks. During the course of the simulation, the DMT powers up in the UMTS cell and starts moving at a speed of 30 miles/hr. The DMT follows a predefined trajectory (as shown in the Fig. 4). The mobility of the DMT causes it to make handoffs from the *UMTS to the WLAN Network* (upon entering the WLAN network) and vice versa (upon exiting from the WLAN coverage area). The UMTS and WLAN networks are integrated at the GGSN node. The WLAN network is considered as a separate packet data network. However, the WLAN coverage area is completely within the UMTS coverage area.

When the DMT powers on, it first completes *GPRS Attach signaling procedure* to establish a PS signaling with the SGSN. This enables it to gain access to the UMTS core services through the UMTS radio stack. Next, the DMT activates a *Packet Data Protocol (PDP) context* in the GGSN. A PDP context is a logical association between a mobile node and a 3G network. The PDP context defines aspects such as QoS, security, billing *etc.* Upon receipt of the Activate PDP Context Request, the SGSN sends a Radio Access Bearer (RAB) Assignment Request to the RNC along with the QoS requested. The UTRAN performs admission control to determine if the request can be granted. If the uplink and downlink have sufficient capacity to accommodate the request, the request is granted. If the request can be granted, the RNC sends a Radio Bearer Setup request to the DMT. On receipt of the Radio Bearer Setup request, the DMT sets up the channel as specified in the request and sends a Radio Bearer Complete to the RNC. When the RNC receives the Radio Bearer Complete, it sends a RAB Assignment Response, which includes the granted QoS, to the SGSN/GGSN. The DMT can send packets to the destination on receipt of the Activate PDP Context Accept message from the SGSN. Before reaching their destination, these packets are first tunneled through the serving RNC and GGSN (using the GTP protocol), and then routed through the IP cloud to their destination node.

The DMT then moves while transmitting and receiving packets using its UMTS in-

terface. These packets are routed through the UMTS core network. At this time, the DMT's WLAN interface is in the passive scan mode and monitors the physical layer for WLAN beacon signals. Upon entering the WLAN coverage area, the WLAN interface receives beacon signals from the WLAN access point. This requires the DMT to initiate a handoff in order to switch to the WLAN access interface. The *criterion for interface selection* used is the bandwidth of the available access network. We have designed and implemented a Switching Module, which is really a software layer responsible for making the access interface switching decisions, and intelligently selecting an access network in the presence of overlapped coverage. The design of this Switching Module and the handoff signaling procedures are explained in the following sections. The GGSN maintains a table mapping of each subscriber's International Mobile Subscriber Identifier (IMSI) and IP Address. As the user moves across different networks, it keeps sending update messages to the GGSN about the change in its IP Address.

5.1 Handoff Process between UMTS and WLAN Networks

When the DMT enters the WLAN coverage area and receives a beacon signal from the WLAN AP, the WLAN_MAC layer computes the signal strength of the received signal. If the strength of the received signal is greater than the minimum packet-reception power threshold value, then the WLAN_MAC layer sends a message (which includes the BSSID [13]) to the Switching Module, indicating that a WLAN AP has been located. This implies that when the signal from the WLAN access point is not stable enough (or has a value lower than the minimum packet-reception power threshold), the handoff to the WLAN is not initiated. Thus, the case of having a weak radio signal at the boundary is accounted for. However, for a transition from the WLAN to a UMTS network, the detection of a weak WLAN signal leads to the handoff to the UMTS network since the UMTS network is considered as an always-on overlay network having a wider coverage area encompassing the WLAN.

If the BSSIDs match, the Switching Module sends a message to the WLAN_MAC layer to begin association with the access-point. Subsequently, the DMT attaches to the WLAN AP. The DMT continues to use both of its radio interfaces for some time: the UMTS interface continues data transmission and reception and the WLAN interface continues to send and receive signaling messages to associate with an access point. This technique is called "soft handoff" and avoids any disruption of service while switching from the UMTS to the WLAN interface. Once the DMT is completely associated with the WLAN AP, the Switching Module sends an Update Message to the GGSN node with its new IP address. Thereafter, the incoming packets are routed by the GGSN to the WLAN access point and delivered to the DMT through the WLAN interface.

When the DMT moves out of coverage of the WLAN network, the WLAN interface detects the reduced strength of the received signal and passes this information to the Switching Module. The Switching Module then sends a message to the GGSN with its UMTS IP address. Thereafter, the packets are tunneled and routed by the GGSN to the RNC, which routes them to the DMT's UMTS interface. A short handoff delay is observed when the handoff is from the WLAN to the UMTS network. This process called "hard handover" occurs when the old connection needs to be released before establishing a new connection.

The reason for using soft handoff in the UMTS-to-WLAN switching and hard handoff in the WLAN-to-UMTS switching is because the network topology is created in such a way that the WLAN network is completely within the UMTS network. So, any mobile device that is within the WLAN network is technically also within the UMTS network. Therefore, soft handoff can be used while moving from UMTS to WLAN network. In this case, the mobile device can continue to use the UMTS interface until a connection is completely established through the WLAN interface. However, while moving from WLAN network into the UMTS network, the mobile device loses its connection from the WLAN network as it moves out of the WLAN coverage area. There is a short time delay between the loss of the WLAN connection and the establishment of the UMTS connection. In this case, the mobile device cannot use the WLAN connection until a UMTS connection is established since the mobile device moves out of the WLAN area before the UMTS connection is fully established. Thus it is not possible to use soft handoff in this case and hard handoff needs to be implemented.

5.2 Performance Evaluation

We evaluated the performance of our dual-mode radio access protocol design and the Switching Module through experimental simulations using various types of applications including: Voice over IP (VoIP) in GSM encoded format, FTP, and HTTP (web browsing). The parameters measured include end-to-end packet delay, file upload time, HTTP page response time, and handoff delay when moving from one network to another. FTP flows are configured with a constant file size of 85000 Kilobytes. Real-time VoIP flows were configured with a constant packet size of 33 bytes. Page size used for HTTP transfers was 3000 bytes.

Table 4 shows the response times obtained in the UMTS and the WLAN networks. In the UMTS coverage area, the FTP upload time is on an average 20.071 seconds and as the user enters into the WLAN coverage and switches transmission through the WLAN interface, the upload response time drops to an average of 0.6 seconds. We obtained similar reductions in the case of end-to-end delays obtained for voice (average of 0.348 sec in UMTS network and 0.005 sec in WLAN network) and HTTP page retrieval time (average of 1.268 seconds in the UMTS network and 0.032 seconds in the WLAN network). The drop in the response times is most likely due to the higher bandwidth offered by the WLAN network (11 Mbits/sec) compared to 144 Kbits/sec offered by the UMTS network.

Table 4. Response time (seconds) for user applications in UMTS and WLAN networks.

Applications	Minimum	Maximum	Average	Network
HTTP	0.909	1.755	1.268	UMTS
FTP	7.094	71.347	20.071	
Voice	0.346	0.349	0.348	
HTTP	0.013	0.049	0.032	WLAN
FTP	0.775	3.657	0.510	
Voice	0.00048	0.00055	0.0005	

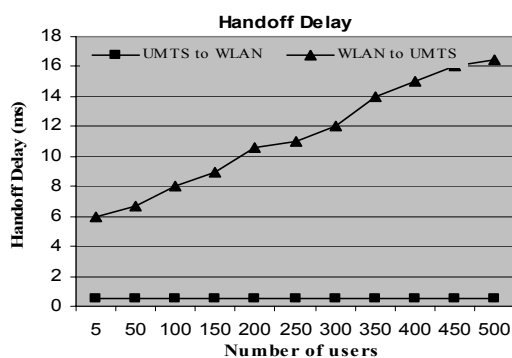


Fig. 5. Handoff delay.

Fig. 5 shows the handoff delay experienced when moving across the UMTS and WLAN networks. The handoff delay values show that there is very little delay when switching the access interface while moving from UMTS to WLAN. This is because of the implementation of soft handover in which data transmission continues to take place through the UMTS interface until the mobile node is completely associated with the WLAN AP and has received a registration confirmation from the GGSN node. Then the Switching Module directs traffic through the WLAN interface. Since the WLAN network is completely within a UMTS cell, it is possible to achieve a smooth handoff when a user moves into a WLAN coverage area. However, when a user moves out of the WLAN coverage area, the WLAN interface detects a weak beacon and informs the Switching Module. This causes the Switching Module to send an IP Address update message to the GGSN. Since it takes some time for the GGSN to receive this message and update the current IP Address of the DMT, a small handoff delay is experienced. However, traffic on the UMTS uplink channel still experiences no delay since the UMTS interface is active and can be immediately used to transmit data. Packets received from the Internet are slightly delayed. However, as observed from Fig. 5 this delay is extremely small and does not introduce a large service disruption period. Our proposed and implemented design architecture that enables seamless internetworking between UMTS and WLAN networks has several benefits. First, the WLAN can be deployed as an independent network and combined ownership of the networks is not required. Second, the approach does not require any modifications to existing 802.11 access points. Third, since all traffic passes through the central GGSN node, it is also possible to use common billing and authentication mechanisms. The low handoff involved when moving across these networks makes roaming completely transparent and seamless to the mobile users and is thus well suited to support real-time applications.

6. A COMPARISON OF 3G-WLAN INTERNETWORKING APPROACHES

Various approaches for integrating WLANs and 3G networks have been proposed in the literature. Some of these approaches include Mobile-IP [5], Gateway [10], Emulator [5], Tight coupling, Loose coupling [9], *etc.* However, each of these approaches integrates the WLAN and 3G networks using the integration point as either SGSN or GGSN.

For example, Mobile-IP and Gateway approaches are based on the GGSN-integration (using the Gi interface). The emulator approach is based on the SGSN integration (using the Gb interface). Furthermore, Tight coupling is a SGSN-based approach where as Loose coupling is a GGSN-based integration scheme. Other characteristics of these approaches have been outlined in Table 1. Consequently, integrating 3G networks with WLANs requires a design that either uses SGSN or the GGSN as the integration point.

Previous research efforts (Tsao *et al.* [5] *etc.*) have analyzed the performance of 3G-WLAN integration approaches with respect to the handoff latency experienced by users while moving from one network to another. In contrast to such previous efforts, the main contributions of this work are as follows:

- We have analyzed the two basic integration architectures (SGSN-based versus GGSN-based) from a design perspective. We evaluated the performance of these integration approaches with respect to the end-to-end packet latency experienced when user traffic from WLAN and UMTS networks is routed through the SGSN (Gi interface) and the GGSN (Gb interface). The latency experienced by the packets in each case is influenced by the operations performed on the different types of WLAN access points connecting the SGSN and the GGSNs (as discussed in section 4.2). Our results (Fig. 3) show that the end-to-end packet latency is lower when GGSN is used as the integration point as compared to the SGSN-integration due to the extra processing overheads involved at the specialized WLAN access point used in the latter case.
- We have proposed a handoff technique that achieves a lower handoff delay as compared to existing GGSN-based handoff approaches (Mobile IP, Gateway, *etc.*) while switching network access from 3G (UMTS) to WLANs. Our approach achieves a seamless network transition between cellular and wireless local area networks by exploiting soft handoff techniques. Fig. 6 shows a comparison of the handoff delays obtained using different types of 3G-WLAN internetworking approaches. The proposed handoff approach yielded the lowest handoff delay thereby demonstrating better performance compared to previously proposed handoff techniques.
- We have discussed the design and implementation of a software layer called the *Switching Module* that enables dual-mode user terminals to perform seamless handovers across cellular and Wireless Local Area Networks using the proposed handoff technique.

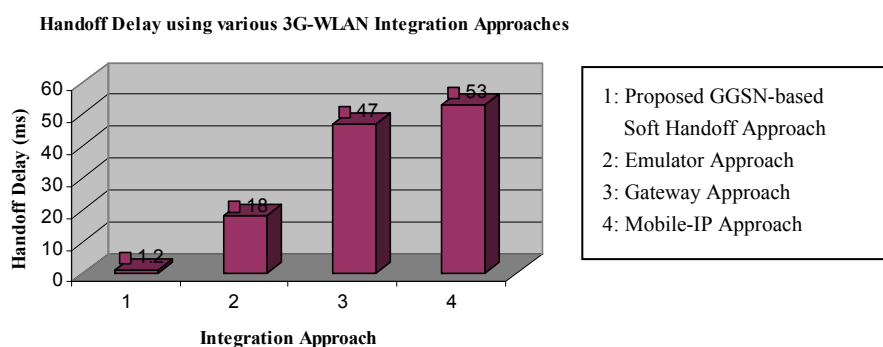


Fig. 6. Handoff delay using various 3G-WLAN integration approaches (UMTS to WLAN).

7. CONCLUSION

Integrated WLAN and UMTS networks benefit users with both high-speed connectivity as well as widespread coverage. Development of architectures that allow interoperability and internetworking between these technologies along with seamless roaming facility is a challenge today. In this work, we have proposed, implemented and evaluated an architecture for a dual interface mobile node that implements a dynamic interface switching algorithm based upon the received signal strength of the WLAN beacon signal. The technique enables smooth handoff process to take place while moving across heterogeneous networks such as UMTS and WLANs. The performance evaluation of the architecture shows minimal handoff delay while switching from UMTS to WLAN networks and a slightly higher handoff delay incurred when moving out of the WLAN coverage and switching radio access to the UMTS interface. The small delays obtained make this architecture feasible for use in both real-time as well as non real-time environments. Another feature of our design approach is that it is based on simple IP services and does not require the use of Mobile IP or IPv6 mechanisms for seamless connectivity.

ACKNOWLEDGEMENT

We thank the anonymous reviewers for their comments that helped us to improve the presentation of this paper. We are also grateful to Dr. Ten-Hwang Lai for his encouragements and support.

REFERENCES

1. IEEE 802.11b, Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications-Higher-Speed Physical Layer Extension in the 2.4 GHz Band, 1999.
2. H. Luo, Z. Jiang, B. J. Kim, N. K. Shankar, and P. Henry, "Internet roaming: a WLAN/3G integration system for enterprises," AT&T Labs-Research, <http://www.nd.edu/~hluo/publications/SPIE02.pdf>.
3. J. D. Vriendt, P. Laine, C. Lerouge, and X. Xiaofeng, "Mobile network evolution: a revolution on the move," *IEEE Communications Magazine*, Vol. 40, 2002, pp. 104-111.
4. TS 125 422-V6.0.0, "Universal mobile telecommunications system," http://hwebapp.etsi.org/action%5CPU/20040120/ts_125422v060000p.pdf.
5. S. Tsao and C. Lin, "Design and evaluation of UMTS-WLAN internetworking strategies," in *Proceedings of IEEE Vehicular Technology Conference*, Vol. 2, 2002, pp. 777-781.
6. M. Buddhikot, G. Chandranmenon, S. Han, Y. Lee, and S. Miller, "Design and implementation of a WLAN/CDMA2000 internetworking architecture," *IEEE Communications Magazine*, 2003, pp. 90-100.
7. B. Hongyang, H. Chen, and L. Jiang, "Performance analysis of vertical handoff in a UMTS-WLAN integrated network," in *Proceedings of the 14th IEEE Conference on Personal, Indoor and Mobile Radio Communications*, Vol. 1, 2003, pp. 187-191.

8. P. Matusz, P. Machan, and J. Wozniak, "Analysis of profitability of inter-system handoffs between IEEE 802.11b and UMTS," in *Proceedings of the 28th Annual IEEE International Conference on Local Computer Networks*, 2003, pp. 210-217.
9. A. Salkintiz, C. Fors, and R. Pazhyannur, "WLAN-GPRS integration for next-generation mobile data networks," *IEEE Wireless Communications*, Vol. 9, 2002, pp. 112-124.
10. S. Tsao and C. Lin, "VSGN: a gateway approach to interconnect UMTS/WLAN networks," in *Proceedings of the 13th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Vol. 1, 2002, pp. 275-279.
11. OPNET Inc., <http://www.opnet.com>.
12. V. Varma, S. Ramesh, M. Barton, G. Hayward, and J. A. Friedhoffer, "Mobility management in UMTS/WLAN networks," *IEEE International Conference on Communications*, Vol. 2, 2003, pp. 1048-1053.
13. S. Uskela, "Key concepts for evolution toward beyond 3G networks," *IEEE Wireless Communications*, Vol. 10, 2003, pp. 43-48.

F. Siddiqui received a M.S. degree in Computer Science from Wayne State University, Detroit, Michigan, and a B.E. degree in Computer Science from Osmania University, India. Currently she is a Ph.D. candidate in the Department of Computer Science at Wayne State University. Her research interests include wireless networks, mobile computing, voice over IP, and network security.



S. Zeadally received the B.A. and M.A. degrees in Computer Science from University of Cambridge, England, and his doctoral degree in Computer Science from University of Buckingham, England, in 1996. He is currently an Associate Professor in the Department of Computer Science and Information Technology at the University of the District of Columbia, Washington DC where he directs the network systems laboratory. He currently serves on the editorial boards of several international journals and he is the Co-Editor-in-Chief of the International Journal of Internet Protocol Technology. He is a Fellow of the British Computer Society (FBCS) and a Fellow of the Institution of Electrical Engineers (FIEE). His research interests include computer networks (wired, wireless), mobile computing, network and system security, ubiquitous computing, multimedia networking, and performance evaluation of networks and systems.