QoS-aware Traffic Offloading in 4G/WiFi Multi-RAT Heterogeneous Networks: Opnet-based Simulations and Real Prototyping Implementation

Ahmed S. Ibrahim*, Ahmed M. Darwish, Shady O. Elbassiouny, and Hani Elgebaly

Intel Corporation, Wireless Communications Lab, Middle East Mobile Innovation Center (MEMIC), Cairo, Egypt
* Electronics and Electrical Communications Department, School of Engineering, Cairo University, Egypt
{ahmed.s.ibrahim, ahmedx.m.darwish, shadyx.o.elbassiouny, hani.elgebaly}@intel.com

***Abstract*— Multiple radio access technologies (Multi-RAT) Heterogeneous Networks (HetNets) are considered as one of the recent fundamental ideas in next generation mobile networks. Multi-RAT HetNets aim to increase the network capacity needed to meet the high data demand of mobile users. In Multi-RAT HetNets, data flow can be offloaded to unlicensed bands (e.g., WiFi) to free some of the resources of the 4G licensed band (e.g. LTE or WiMAX). However, it needs to be guaranteed that the WiFi can provide the needed Quality of Service (QoS). In this paper, we propose multiple QoS-aware 4G/WiFi offloading schemes that try to maximize the utilization of the WiFi air interface, taking into consideration the network loading conditions and maintaining the required QoS. The proposed offloading schemes are evaluated in two 4G/WiFi Multi-RAT testbeds. The first testbed is based on the Opnet network simulator, while the second testbed is based on a real small-cell prototype. The performance evaluation of these offloading schemes is shown in a time-cost tradeoff graph.**

Keywords- HetNets; LTE; Multi-RAT; Opnet; Prototype; WiFi; WiMAX.

I. INTRODUCTION

Recently, there has been increased interest in Heterogeneous Networks (HetNets) as a means to meet the huge data demand needed for the current and future mobile applications. For example, various HetNets topics are actively investigated in the 3GPP-Long Term Evolution (LTE) [1] and Worldwide Interoperability for Microwave Access (WiMAX) IEEE 802.16 [2] standards. Research in HetNets can be divided into two main topics, namely, multi-tier and multiple radio access technologies (multi-RAT). The multi-tier HetNets comprise a hierarchical deployment of low power, smaller footprint Femto/Pico/Relay stations within a Macro-cell coverage area, aiming to increase the network capacity and provide reliable indoor coverage [3]. Interference mitigation (IM) is one of the challenging topics in multi-tier HetNets, and we have previously proposed multiple IM techniques that span the time, frequency, space, and power dimensions [4], [5].

The multi-tier HetNets utilizes one air interface only (e.g. LTE). However, multiple air interfaces can be utilized too, which is the case of multi-RAT HetNets [6], [7], [8], [9], [10], [11], [12], [13]. Multi-RAT HetNets allow the support of mobile users with more than one air interface, such as the combination of the licensed 3GPP LTE along with the unlicensed IEEE 802.11 Wireless Local Area Network (WLAN) technologies in order to increase the overall network capacity with reduced cost. It is known that the 4G technology (LTE or WiMAX) utilizes license bands with more charging cost compared to utilizing the unlicensed WiFi air interface. Hence, it is cost-effective to be utilizing the WiFi interface, as long as it is providing the required service quality.

In the literature, there has been a number of works considering 3G-based Multi-RAT systems such as [6], [7], [8]. A feasibility study on multiple radio access (MRA) networking was introduced in [7] and an implementable MRA was proposed in [8]. Beyond 3G, Multi-RAT systems were analyzed with the objective of network capacity maximization in [9], [10], [11], [12]. Other than capacity-maximization, we have focused on minimizing the transmission power in [13], and we have proposed multiple minimum-power LTE/WiFi offloading schemes that minimize the transmission power, while guaranteeing the required QoS.

In this paper, we aim to propose multiple QoS-aware 4G/WiFi offloading schemes that try to maximize the utilization of the WiFi interface, while maintaining the required QoS. Maximizing the WiFi utilization, compared to the 4G utilization, corresponds to lowering the transmission cost. In order to propose these offloading schemes, we develop two independent testbeds. The first testbed is based on the Opnet network simulator [14], in which the available 4G technology is WiMAX, while the second testbed is based on a real prototype, in which the available 4G technology is LTE. We point out that considering more than one 4G technology (LTE or WiMAX) in more than one testbed (simulations or real prototype) strengthens this paper and generalizes the presented ideas.

The main contributions of this paper are as follows. First, we develop an Opnet-based WiMAX/WiFi multi-RAT HetNet simulation environment, which consists mainly of a multi-RAT base station (BS) or small cell and a *newly* created multi-RAT UE. Second, we propose a QoS-aware offloading scheme, which assigns the users to either the WiMAX or the WiFi air interface, depending on the network loading condition. The proposed scheme utilizes the WiFi interface unless it cannot guarantee the required QoS. Third, we build a real LTE/WiFi Multi-RAT small-cell prototype and explain how this prototype can be utilized in sending a video clip over the two air interfaces (LTE and WiFi). Finally, we propose two QoS-aware offloading schemes, implement them on the small-cell prototype, and evaluate their time-cost performance tradeoff graph.

The rest of the paper is organized as follows. In the next section, we focus on the Opnet-based Multi-RAT simulation environment, clarify the problem, and evaluate the proposed

 ISBN: 978-1-61208-284-4

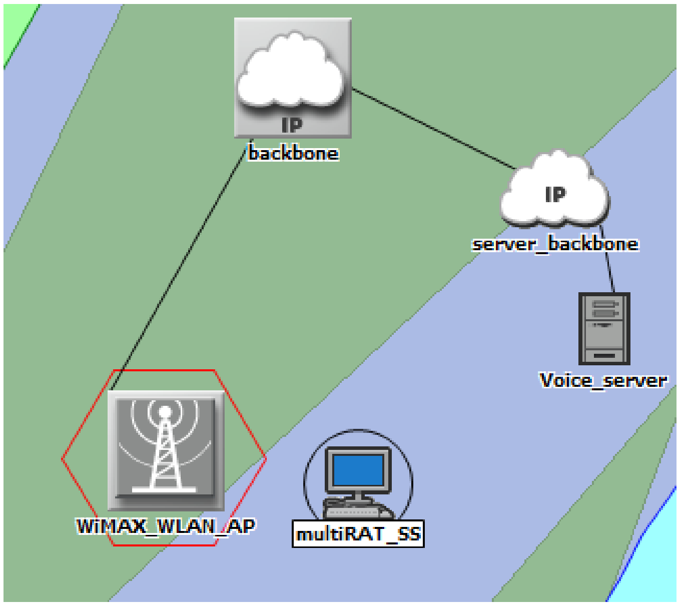


Fig. 1. Opnet-based Multi-RAT HetNet.

offloading scheme. The real LTE/WiFi small-cell prototype is introduced in Section III, after which two additional offloading schemes are proposed. Finally, Section IV concludes the paper.

II. Opnet-based QoS-aware WiMAX/WiFi Offloading Scheme

In this section, first we introduce our developed Opnet-based WiMAX/WiFi Multi-RAT HetNet environment. Second, we explain our motivation to take into consideration the QoS while designing the offloading schemes. Third, we pose the problem statement. Finally, we propose a QoS-aware WiMAX/WiFi offloading scheme.

A. Opnet-based Multi-RAT Heterogeneous Network Model

Fig. 1 depicts a multi-RAT HetNet, which consists of multi-RAT small cell (or BS), multi-RAT UE, VoIP server, and core network (server backbone and IP backbone). The multi-RAT BS along with the VoIP server and the core network are already built in the Opnet library (v14.5). However, there was no Multi-RAT UE that can communicate with both WiFi and WiMAX, which represented our first challenge.

Our first contribution is to create a novel multi-RAT (WiMAX/WiFi) UE that can receive the data over either the WiFi or the WiMAX air interfaces. Fig. 2 depicts the node model of the proposed Multi-RAT UE in details. As shown, the set of all the upper layers (e.g. IP) remains untouched as in the single-RAT node model. On the other hand, the lower layers, namely, Address Resolution Protocol (ARP) and Media Access Control (MAC) are duplicated for each interface. The MAC module is the one differentiating WiMAX from WiFi. We note that each UE will have one specific IP address, through which it can be addressed in the network. This IP address will not be changed if the UE is utilizing either the WiMAX or the WiFi air interface. Moreover, either the small cell or the UE may decide either to activate the WiMAX or the WiFi air interfaces.

TABLE I
WIMAX AND WIFI AIR INTERFACE CONFIGURATION PARAMETERS.

Parameter	Value
WiMAX PHY profile	OFDMA
WiMAX Bandwidth	20MHz
WiFi PHY profile	Extended rate PHY (802.11g)
WiFi Data rate	11 Mbps
Multipath channel model	ITU Pedestrian A
Pathloss model	Free space

In order to test the performance of the developed HetNet environment, we assume the UE initiates a VoIP call with the VoIP server. We note that we focus on real time application (mainly VoIP) and evaluate its QoS, measured in terms of Mean Opinion Score (MOS). The VoIP data will be sent over WiFi or WiMAX to the small cell, which in turn will be delivered through the core network to the VoIP server. The VoIP packets from the VoIP server to the UE follow the same path in reversible order. Table I shows the major configuration parameters for both WiMAX and WiFi air interfaces.

Fig. 3 depicts the performance of the VoIP application, measured in MOS, for both the WiMAX and WiFi cases. In addition, it shows the loading condition on both air interfaces, to indicate which interface is activated. First, we consider the case of having WiMAX as the active interface, which is depicted in Fig. 3 (a). As shown, the load on the WiMAX network is non-zero, while it is zero for the WiFi case. The MOS of the VoIP application is around 3.8, which gives great user experience. Second, the WiFi is activated and the resulting performance is shown in Fig. 3 (b). Similar to the WiMAX, we find that the WiFi also achieves the same MOS value.

From Fig. 3, we conclude that if the VoIP application is the only traffic in the network, offloading to WiFi does not degrade the MOS below that achieved by the WiMAX. In such scenario, traffic offloading from WiMAX to WiFi should happen immediately once WiFi is available. However, this might not be necessarily true in case there is other additional traffic in the network. This case will be discussed in the next sub-section.

B. Motivation and Problem Statement

In general, we note that there is a major difference between the WiMAX (IEEE 802.16e) and the WiFi (IEEE 802.11g), which is the consideration of the QoS classes. Particularly, in WiMAX each type of application is assigned a specific class, which corresponds to the needed QoS. Traffic types with high QoS class (e.g. VoIP) are assigned resource blocks with high priority, compared to traffic types with lower QoS classes (e.g. best effort as in web browsing). On the contrary, in WiFi all the traffic types are treated equally in the best effort class. This fact follows mainly from the utilized multiple access scheme,

ISBN: 978-1-61208-284-4

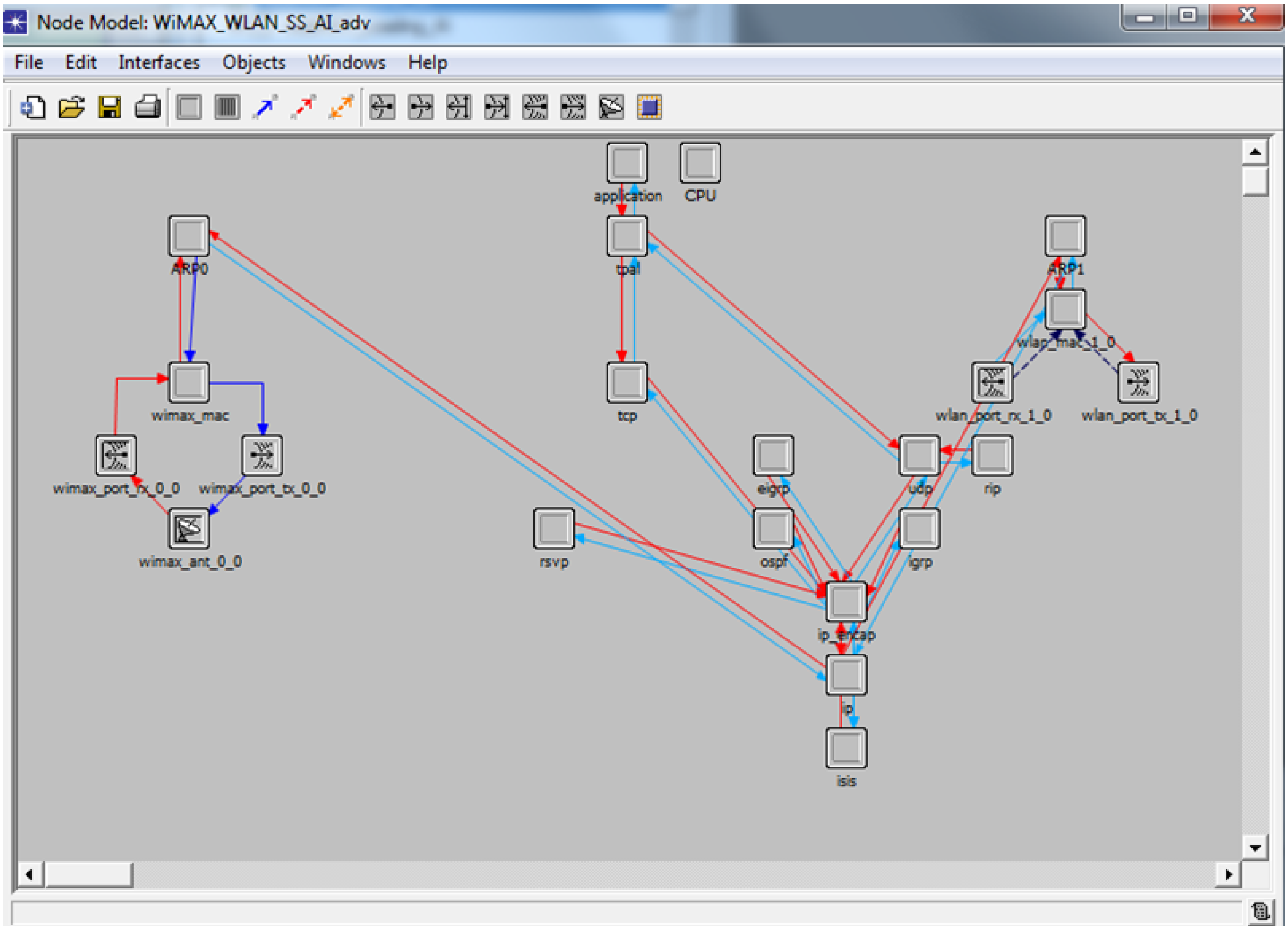


Fig. 2. Newly created Opnet Multi-RAT UE.

namely, Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA). Hence, in WiFi there is no distinction between the real time applications and non real time ones. Such issue may degrade the performance of the VoIP applications sent over WiFi, once the network gets congested.

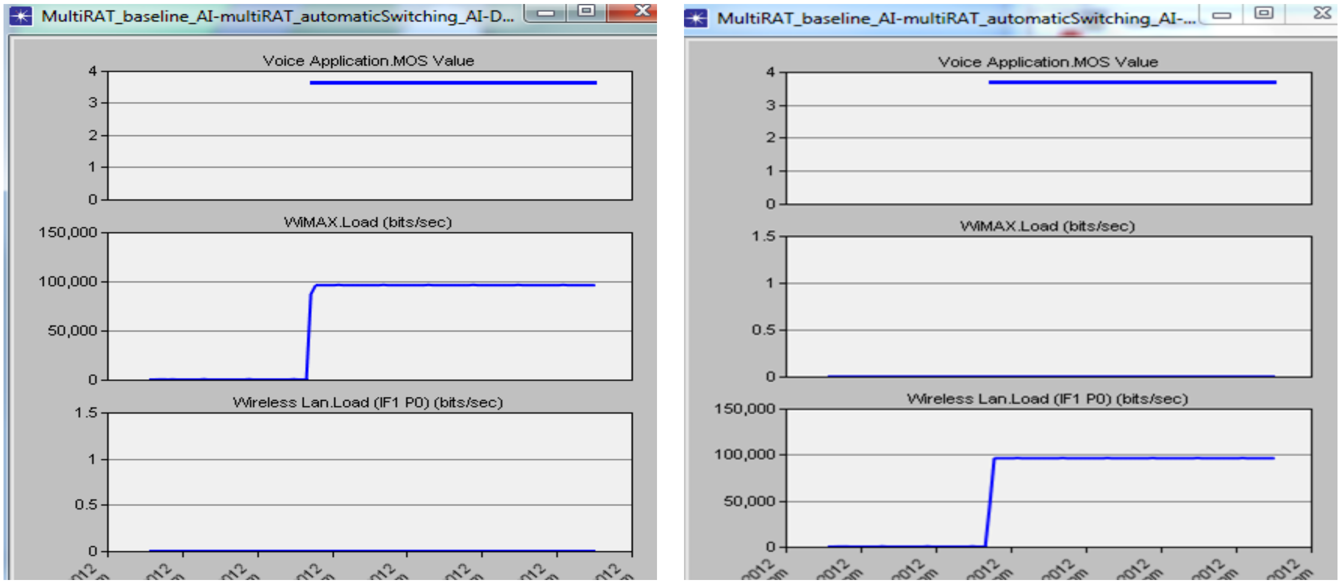
In order to investigate this issue, we revisit the multi-RAT scenario shown in Fig. 1 and apply a best effort background data traffic of rate 2 Mbps, in addition to the existing VoIP application. This background data traffic may represent the accumulation of many UEs accessing the same air interface, or it may represent data traffic to/generated from the same UE. Fig. 4 shows the network performance in the case of WiMAX and WiFi interfaces for 2 Mbps background best-effort data traffic.

As shown in Fig. 4 (a), the WiMAX air interface maintains the high value of the VoIP MOS. This high performance is a result of assigning high QoS class to the VoIP application. Hence, the WiMAX resource blocks needed for the VoIP application are always allocated irrespective of the other background traffic. On the contrary, we find in Fig. 4 (b) that the WiFi air interface dramatically degrades the MOS of the VoIP application. Such performance degradation is because the best effort background traffic is competing with the VoIP application, which results in lower chances for the VoIP application in accessing the WiFi air interface.

From Fig. 4, we conclude that offloading from WiMAX (or 4G in general) to WiFi should not happen automatically once WiFi access is available. Other factors (e.g. background traffic) need to be considered to make such offloading decision. This finding represents our motivation to develop a QoS-aware offloading scheme that takes into consideration the loading condition of the network. Consequently, our probem can be formulated as follows: Given a multi-RAT (4G/WiFi) system, find an offloading scheme that utilizes the WiFi interface the most (i.e. minimizes the charging cost), while guaranteeing the required QoS. The solution to this problem will be proposed in the next sub-section.

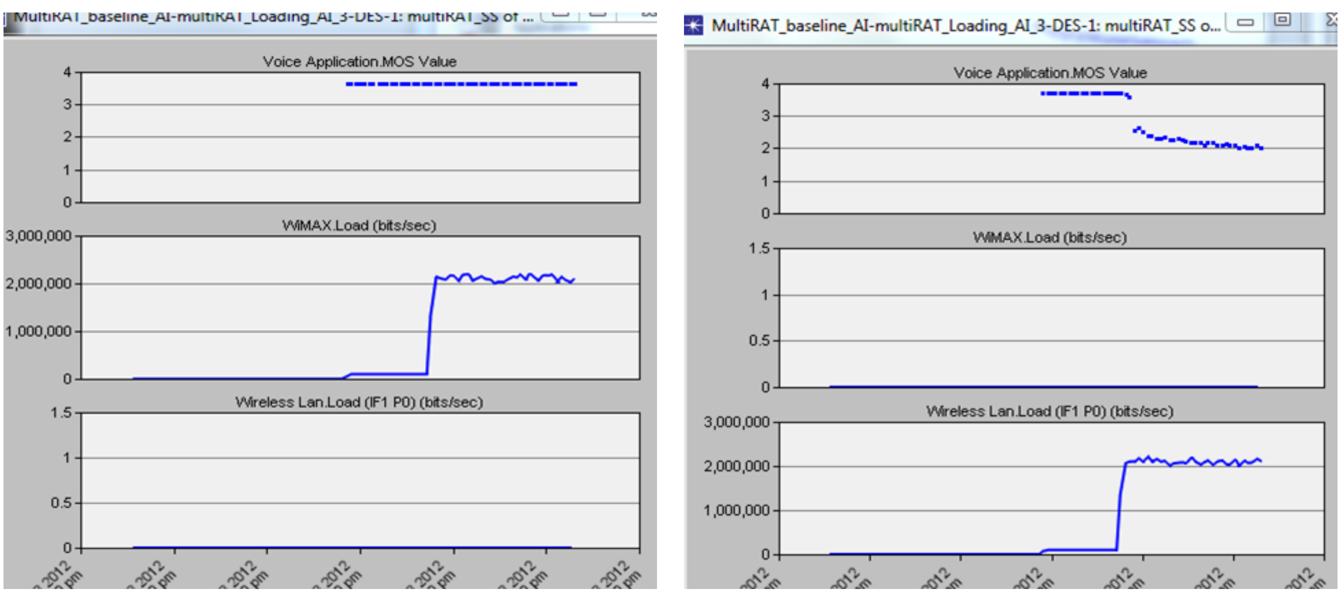
C. QoS-aware Traffic Offloading Scheme in WiMAX/WiFi Multi-RAT HetNets

Fig. 5 depicts the flow chart of the proposed QoS-aware offloading scheme, and it can be described as follows. Initially, the 2 RATs are deactivated. Then, the UE senses the total data rate (loading condition) on each interface. If the loading



(a) WiMAX On (b) WiFi On

Fig. 3. Performance of WiMAX and WiFi interfaces for VoIP-only network.



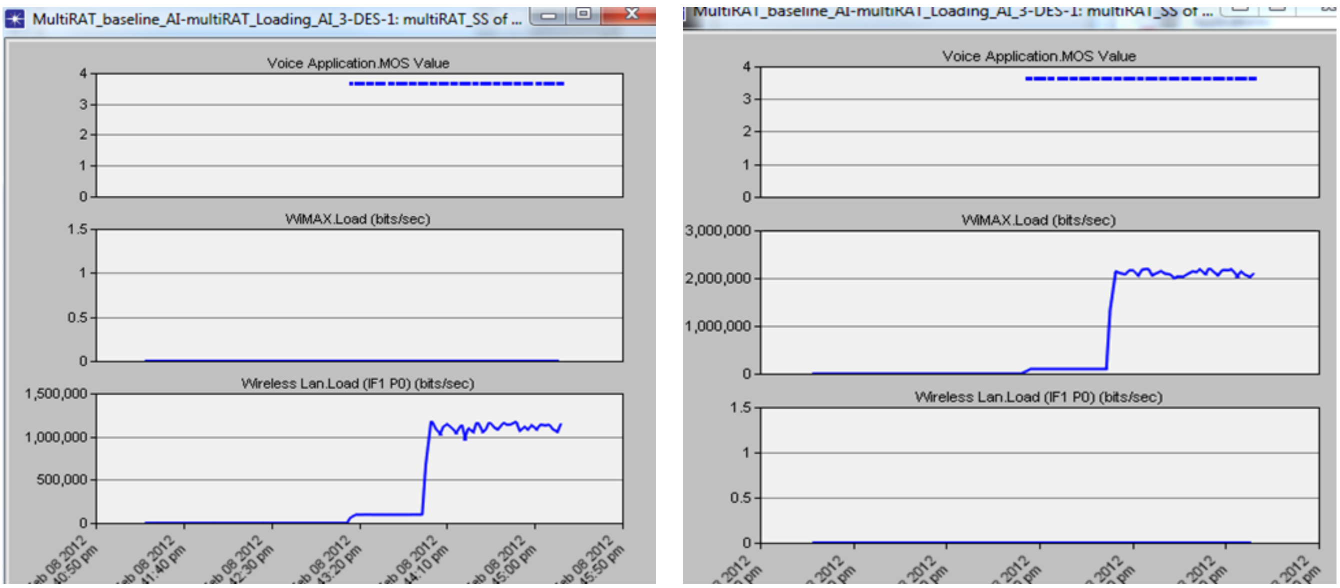
(a) WiMAX On (b) WiFi On

Fig. 4. Network performance with 2Mbps background best-effort traffic.

rate is higher than a specific threshold, then it activates the WiMAX interface. In this case, the WiFi will not be able to guarantee the required QoS of the real-time application. Otherwise, it offloads the traffic to the WiFi interface. Utilizing our developed Opnet-based HetNet environment, we have investigated the threshold after which the WiFi breaks down and starts downgrading the VoIP MOS, and found that to be 1.5 Mbps. In other words, for background traffic higher than 1.5 Mbps, the VoIP MOS degrades as was shown in Fig. 4 (b).

We have implemented the proposed offloading scheme in the Opnet-based Multi-RAT environment at the IP process module of the Multi-RAT UE node, shown in Fig. 2. We have tested the offloading scheme by applying different background traffic rates, which are 1 Mbps and 2 Mbps. Fig. 6 shows the performance of the offloading scheme at these two scenarios. As shown, for 1 Mbps loading scenario in Fig. 6 (a), the WiFi interface is activated and the MOS is maintained at high value. In this case, offloading to WiFi guarantees the required QoS. As for the higher loading rate (2 Mbps in Fig. 6 (b)), the WiMAX is activated and the high MOS is maintained. Hence, our proposed QoS-aware offloading scheme achieves the required QoS with minimum charging cost. More precisely, it offloads the traffic to the unlicensed minimum-cost WiFi

 ISBN: 978-1-61208-284-4



(a) WiFi On at 1 Mbps loading

(b) WiMAX On at 2Mbps loading

Fig. 6. Network performance with QoS-aware offloading scheme at different network loading conditions.

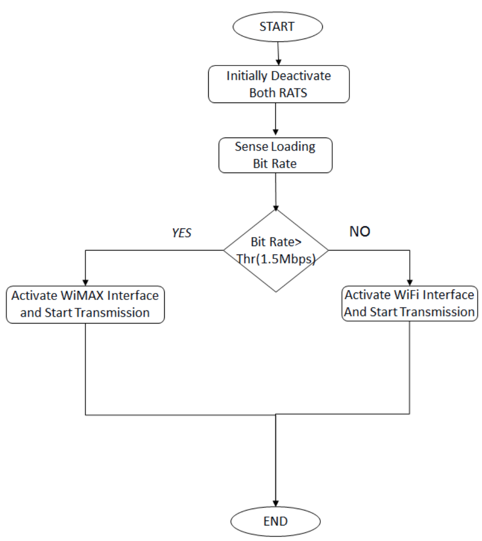


Fig. 5. Flow chart of the QoS-aware traffic offloading scheme.

only if the required QoS is guaranteed.

III. Real Prototype-based QoS-aware LTE/WiFi Offloading Schemes

In the previous section, we have focused on Opnet-based multi-RAT simulation environment. In this section, we present a real prototype that represents the implementation of LTE/WiFi small cell. Such prototype is extremely necessary to validate the concept of multi-RAT system and to augment our simulation testbed with a real one. In this prototype, we utilize Intel-proprietary LTE software stack, consisting of all the LTE layers (RRC/PDCP/RLC/MAC). As for the WiFi, we utilize the commercial Intel 6205 WiFi card, which can be set to work in the access point mode. Hence, we create an integrated LTE/WiFi small cell that has both WiFi and LTE software stacks.

A. Real LTE/WiFi Multi-RAT Prototype

Fig. 7 depicts the system architecture of the implemented multi-RAT prototype. The prototype consists of two terminals: the first terminal has an integrated LTE/WiFi small cell, and the second terminal represents the UE side. The LTE data between the two terminals is physically sent over the Ethernet cable, which mimics the LTE PHY layer, while the WiFi data is physically sent over the WiFi air interface. At the small cell side, the LTE down-link path is represented by the (PDCP-TX/RLC-TX/MAC-TX) path, and the LTE up-link path is represented by the (MAC-RX/RLC-RX/PDCP-RX) path. The WiFi path is represented by the (NDIS/MAC/PHY) path. At the UE side, we create a simple software client that can send/receive data to/from the Ethernet and WiFi interfaces.

A photo of the real prototype is shown in Fig. 8. The small-cell and the UE terminals are shown. In addition, we have created a WiFi loading environment, which consists of two terminals having file transfer between each other. Such file transfer causes huge degradation on the WiFi link, which results in huge WiFi congestion.

In order to test this prototype, we have created a demonstration experiment in which a video clip is sent from the small-cell terminal to the UE terminal via both LTE and WiFi software stacks. The Ethernet link is used to transfer the LTE data, while the WiFi is sent over the conventional WiFi air interface. Initially, a part of a video clip (a number of packets) is sent over the LTE. Once the WiFi coverage is available, the transmission is offloaded to the WiFi. Once the WiFi

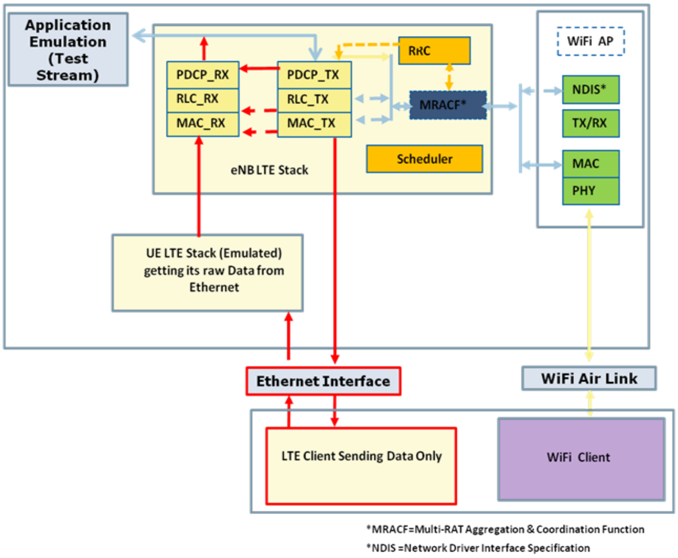


Fig. 7. System architecture of the LTE/WiFi Multi-RAT prototype.

Fig. 8. Real setup of the LTE/WiFi Multi-RAT prototype.

network is congested (due to loading the WiFi link with the file transfer), the rest of the video transmission is sent over LTE. Such offloading scheme was implemented in the multi-RAT aggregation and coordination function (MRACF) at the small cell side. At the receiver side, the received packets over the two interfaces are assembled, and the video is displayed.

The small-cell prototype described above is utilized in evaluating the performance of our proposed QoS-aware offloading schemes. Particularly, we consider two performance criteria to evaluate these schemes, namely, the time and the cost. First, the time metric represents the total number of transmission slots needed to successfully receive the whole video file. Second, the cost metric represents the number of cost units

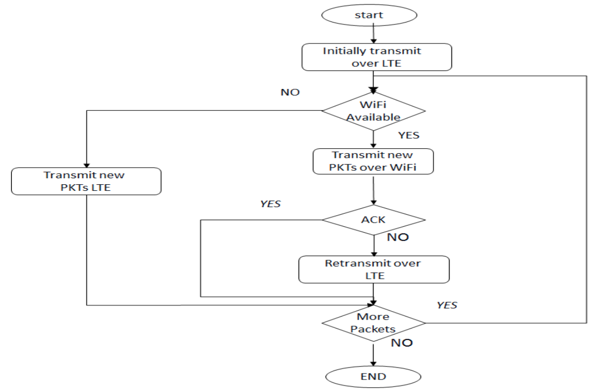


Fig. 9. Flow chart of the “LTE/WiFi Offloading Retransmit-Once” offloading scheme.

needed to successfully receive the video file. We assume that the video file is divided into a number of packets. Each packet requires 1 transmission slot and costs 1 unit over LTE and zero units over WiFi. These cost values reflect the fact that WiFi is unlicensed band, while LTE is a licensed one. The proposed offloading schemes are presented in the next sub-section.

B. QoS-aware LTE/WiFi Offloading Schemes

We propose 2 QoS-aware LTE/WiFi offloading schemes. The first scheme is entitled “LTE/WiFi Offloading Retransmit-Once” and its flowchart is depicted in Fig. 9. In this scheme, the packets are initially sent over LTE. Once the WiFi coverage is available, the data is offloaded to the WiFi air interface. Due to the created congestion on the WiFi, we have developed an acknowledgment mechanism that allows the UE to transmit back to the small cell either a positive acknowledgment (ACK) or a negative acknowledgment (NACK) for each received packet. A NACK is declared if a particular packet index is missing. Once a NACK is declared, the erroneous packet is retransmitted over the LTE interface. The LTE is utilized to guarantee a congestion-free retransmission. As for the next new packet, the WiFi is activated again and it carries the upcoming packet. Hence, the description in the name “Retransmit-Once.” This procedure continues until the video is successfully received.

The second QoS-aware offloading scheme, namely “LTE/WiFi Offloading Retransmit-Rest”, is similar to the first one in many aspects except in the response following the retransmission of the erroneous packet. More precisely, once an erroneous packet is retransmitted over the LTE, the LTE air interface is kept active and transmits the rest of the video packets. Hence, the second scheme utilizes the LTE interface more and hence it costs more. This is shown in the time-cost tradeoff curve shown in Fig. 10.

Fig. 10 depicts the different offloading schemes results where the x-axis represents the cost of transmission (0 for WiFi packet and 1 for LTE packet) and the y-axis represents the total time (in transmission slots) needed to transmit one video clip of size 23 packets. The “LTE-only” point corresponds to having LTE transmission only and results in the minimum

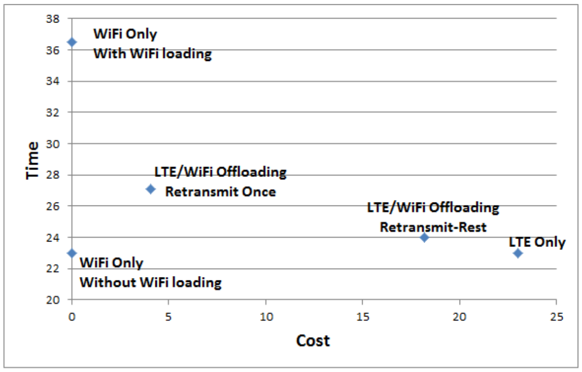


Fig. 10. Time-Cost tradeoff curve for the LTE/WiFi Multi-RAT offloading schemes.

time of 23 units, but with the maximum cost of 23 units. In order to reduce the cost, we have proposed the 2 QoS-aware offloading schemes described earlier. We find that the "LTE/WiFi Offloading Retransmit-Rest" slightly reduces the cost to 18, while slightly increasing the time to 24 units. The extra time units are due to the retransmission time slots happening because of congestion and packet drop over the WiFi link.

On the other hand, the "LTE/WiFi Offloading Retransmit-Once" reduces the cost much more to 4 units while slightly increasing the time to 27. Compared to "LTE-only" case, we find that the "LTE/WiFi Offloading Retransmit-Once" achieves cost saving of 83% while increasing the time by 15% only. For reference, we have also plotted the points corresponding to the WiFi-only transmission. More specifically, we observe "WiFi-only without WiFi loading," which results in both minimum time and zero cost. If loading is injected, the required time is increased to 37 units, which is shown in the "WiFi-only with WiFi loading."

IV. Conclusion and Future Work

In this paper, we have developed 2 4G/WiFi Multi-RAT testbeds and proposed various QoS-aware offloading schemes. In particular, we have developed an Opnet-based WiMAX/WiFi multi-RAT simulation environment. We have shown, via this simulation platform, that the WiFi degrades the VoIP quality as the background traffic rate exceeds 1.5 Mbps. On the other hand, the WiMAX preserves high quality for the VoIP application irrespective of the background traffic, as it guarantees the needed QoS. Consequently, we have proposed a QoS-aware offloading scheme that decides either to offload the traffic to the WiFi or to utilize the WiMAX, based on the network loading data rate. Such offloading scheme achieves the required rate with minimum charging cost.

In addition to developing the Opnet-based simulation environment, we have also developed a real LTE/WiFi prototype that can be used to transmit any file over the 2 air interfaces. Finally, we have proposed and implemented 2 QoS-aware LTE/WiFi offloading schemes and have evaluated the time-cost performance tradeoff. We have shown that by considering our proposed QoS-aware offloading we can achieve cost saving of 83% while increasing the time by 15% only, compared to "LTE-only" case.

In the future, we will invesitgate which of the different LTE layers (PDCP,RLC,MAC) is the best one at which we can offload the data to the WiFi. We will utilize our developed real LTE/WiFi prototype as a testbed to evaluate the pros and cons of the different offloading points.

V. Acknowledgment

The authors would like to thank Dr. Nageen Himayat, Dr. Kirk Skeba, Dr. Shilpa Talwar, Mohamed Fadel, Wafaa Taie, Mohamed Yousef, and Hossam Samy of Intel Wireless Communications labs, for their constructive discussions and cooperation throughout this work.

References

[1] 3GPP TS36.300 V9.0.0, "Evoloved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (EU-TRAN)," 2009.

[2] IEEE 802.16e, "IEEE standard for local and metropolitan area networks part 16: Air interface for fixed and mobile broadband wireless access systems," 2009.

[3] V. Chandrasekhar, J. G. Andrews, and A. Gatherer, "Femtocell networks: A survey," in *http://arxiv.org/ftp/arxiv/papers/0803/0803.0952.pdf*, 2008.

[4] A. S. Ibrahim, A. Darwish, E. Makled, and H. Elgebaly, "Adaptive interference mitigation techniques for femtocells," *Proc. IEEE Wireless Communication and Networking Conference (WCNC'12)*, Apr. 2012.

[5] E. Makled, A. S. Ibrahim, A. Darwish, and H. Elgebaly, "Non-unanimous power inter-cell interference coordination in heterogeneous networks," *Proc. the IEEE 75th Vehicular Technology Conference Spring (VTC'12-Spring)*, May 2012.

[6] K. Dimou, R. Agero, M. Bortnik, R. Karimi, G. Koudouridis, S. Kaminski, H. Lederer, and J. Sachs, "Generic link layer: a solution for multi-radio transmission diversity in communication networks beyond 3g," in *Proc. IEEE Veh. Techn. Conf. (VTC)*, Sept. 2005, vol. 3, pp. 1672–1676.

[7] G. P. Koudouridis, R. Agero, E. Alexandri, M. Berg, A. Bria, J. Gebert, L. Jorguseski, H. R. Karimi, I. Karla, P. Karlsson, J. Lundsj, P. Magnusson, F. Meago, M. Prytz, and J. Sachs, "Feasibility studies and architecture for multi-radio access in ambient networks," in *Wireless World Research Forum (WWRF) Meeting*, Dec. 2005.

[8] E. Gustafsson and A. Jonsson, "Always best connected," *Wireless communication, IEEE Trans*, vol. 10, no. 1, pp. 49–55, Feb. 2003.

[9] Y. Choi, H. Kim, S. Han, and Y. Han, "Joint resource allocation for parallel multi-radio access in heterogeneous wireless networks," vol. 9, no. 11, pp. 3324–3329, Nov. 2010.

[10] L. Liu, Y. Wei, Y. Bi, and J. Song, "Cooperative transmission and data scheduling mechanism of multiple interfaces in multi-radio access environment," in *Proc. Global Mobile Congress (GMC)*, Oct. 2009.

[11] S. Ghamari, M. Schellmann, M. Dillinger, and E. Schulz, "An approach for automated spectrum reforming for multiple radio access technologies," in *Proc. Technical Symposium at Telecom World (ITU WT)*, Oct. 2011.

[12] W. Song and W. Zhuang, "Multi-service load sharing for resource management in the cellular/wlan integrated network," vol. 8, no. 2, pp. 725–735, Feb. 2009.

[13] M. Fadel, A. S. Ibrahim, and H. Elgebaly, "Qos-aware multi-RAT resource allocation with minimum transmit power in multiuser OFDM system," *Proc. the IEEE Globecom International Workshop on Emerging Technologies for LTE-Advanced and Beyond-4G (GC'12 Workshop - LTE-B4G)*, Dec. 2012.

[14] "http://www.opnet.com/," *v 14.5*.

 ISBN: 978-1-61208-284-4