Optimized Flow Management using Linear Programming in Future Wireless Networks

Umar Toseef*†, Andreas Timm-Giel*
*Institute of Communication Networks
Hamburg University of Technology
Schwarzenbergstr. 95E, 21073 Hamburg, Germany
Email: {umar.toseef, timm-giel}@tuhh.de

Carmelita Görg†
†TZI ComNets - University of Bremen
Otto-Hahn-Allee, NW1
Bremen, Germany
Email: {umar,cg}@tzi.de

***Abstract*—There have been tremendous advances over the past decades when it comes to wireless access technologies. Nowadays, mobile devices are equipped with several wireless access technologies like 3G, 4G or WiFi. Currently, these mobile devices can communicate using one access technology at a time. However, there is a big potential for improving network capacity and enhancing user 'Quality of Experience' if these access technologies are integrated. Such an integration would allow access technologies to cooperate and work simultaneously in a heterogeneous environment from which both the end users as well as the mobile operators can benefit. In this paper, it is investigated how to tackle the simultaneous usage of wireless access technologies. For this purpose, a practical example of a 3GPP LTE and a non-3GPP WLAN integrated heterogeneous network is considered. Furthermore, a novel decision mechanism is proposed, that focuses on optimizing the flow management of user traffic flows based on a mathematical formulation of the system. The mathematical model is implemented using Linear Programming techniques. The paper demonstrates the gains and benefits that are achieved from using such innovative decision mechanism as well as the benefits that arise from the simultaneous usage of wireless heterogeneous accesses.**

Keywords: ***LTE and WLAN, Resource Allocation, User QoE, Heterogeneous Networks, Linear Programming.***

I. Introduction

There are various prevailing standards of wireless access technologies in the current communication market, such as 3GPP (3rd Generation Partnership Project), non-3GPP, 3GPP2, etc. Admittedly, each type of these access technologies has certain advantages, which justify its existence in this age of evolution of technology. For example, 3GPP networks are more efficient in terms of handling high traffic demands, providing QoS (Quality of Service) guarantees and the extended coverage. Whereas, the non-3GPP technologies like IEEE 802.11 [2] are simple to operate and therefore need less investment and operation & maintenance cost. On the other hand, the wireless portable devices are becoming increasingly popular and it is widely expected that such devices will outnumber any other forms of smart computing and communication in near future. With the capability of connecting through several types of 3GPP and non-3GPP access technologies these devices run a wide variety of bandwidth demanding services including high speed data delivery and multimedia communication. However, due to the limitations of today's network architecture these devices can connect to one access technology a time. It is proposed that in future networks of heterogeneous access technologies, the ever increasing bandwidth demands of the portable devices can be better addressed through the bandwidth resource aggregation of multiple networks. This would create a win-win situation for the network operators and the users. 3GPP standardization has already envisioned the possible benefits from the cooperation of 3GPP and non-3GPP networks and has come up with such integration standards[5]. 3GPP specified System Architecture Evolution (SAE) allows mobile users to roam between 3GPP and non-3GPP access technologies with seamless mobility provided through Proxy Mobile IPv6 (network based mobility) and Dual Stack Mobile IPv6 (host based mobility) [5]. The 3GPP SAE architecture, however, does not support the user multi-homing i.e. simultaneous user connection to more than one access network. In order to investigate the achievable advantages through the support of user multi-homing, the existing 3GPP standards needs extensions. Moreover, the issues related to an efficient management of aggregated bandwidth resource should also be addressed when multi-homing support is realized.

The focus of this work is to extend the 3GPP SAE architecture in realizing multi-homing support for users and propose a solution to make an optimum use of aggregated bandwidth resources and network diversity in a multi-homing scenario.

The rest of the paper is organized as follows: Section II describes how the current 3GPP SAE architecture is extended to provide users with multi-homing support. Section III describes the importance of flow management function in a heterogeneous network, and Section IV explains the linear programming technique used to achieve an optimized flow management operation. Finally, Section V provides the proof of concepts through the discussion of simulation results of an investigated realistic scenario.

II. Network simulation model

This work follows the proposal of 3GPP specifications in the integration of 3GPP access technology (namely LTE) and trusted non-3GPP access technology (namely legacy WLAN 802.11g), where host based mobility solutions, i.e., Dual Stack Mobile IPv6 is considered. For this purpose, a simulation network model has been implemented using the OPNET [6]

 ISBN: 978-1-61208-229-5

network simulator. This includes the detailed implementation of LTE network entities following the 3GPP specifications [8]. Simulation models of WLAN access points as well as, the common protocol layers like application, TCP/UDP, IP, Mobile IP, Ethernet, etc., come from the OPNET standard library [6]. The home agent (HA) function is located at the Packet Data Network (PDN) gateway. The remote server acts as a correspondent node (CN) from where mobile users access application services; (see Fig. 4). A user can have up to two active network interfaces, one for each access technology. Further details about the simulator can be found in [2].

III. Flow Management

Flow management helps a network operator or a user to make use of the two network paths from available access technologies. In general, there are two options of managing traffic flows for a multi-homed user. The first option is to carry one complete application traffic flow over one path of choice, this is known as “traffic flow switching”. For example, a user can decide to keep his TCP based traffic flows on the WLAN access network while his VoIP/video traffic follows its way over the LTE network. The second option is to divide the traffic flow into several smaller sub-flows where each sub-flow is carried over one network path. This will be called “traffic flow splitting”. For example, a user watching HD video streaming of a football match can distribute the video traffic flow over the WLAN and LTE network as long as the user is in the overlapped coverage of both networks. In this work, both of these options will be used based on the requirements of global optimal resource allocation goal.

In a wireless access network, frequency spectrum and its usage time are the main network resources, which are shared by all users. If a user has good channel conditions, he can use higher modulation schemes and achieve higher spectral efficiency. High spectral efficiency allows the user to transmit more data bits for a given amount of network resources. The opposite is true for a user who is suffering from bad channel conditions. The amount of network resources of a wireless network are determined from the designed parameters like frequency spectrum, transmission technology, antenna gains etc. Therefore, each network has a fixed amount of network resources and the network performance itself depends on the fact with what spectral efficiency these resources are utilized. In order to achieve higher data rates a network resource scheduler should select those users who can attain high spectral efficiency, and therefore, need less network resources per unit data rate. In this work, we adapt the term “network path cost” for the required network resources per unit data rate. For a user, its network path cost can be accessed through cross layer information from the MAC layer of the corresponding access technologies. In the following subsections, it is shown how the network path cost can be computed for users in LTE network and how the achievable user throughput can be estimated in WLAN network.

A. *Network path cost for LTE*

LTE performs a managed scheduling of available bandwidth resources. The smallest unit of bandwidth resource is referred to as a physical resource block (PRB) in the LTE specification. Based on the allocated frequency spectrum size LTE has a certain number of PRBs. The LTE MAC scheduler residing at the eNodeB schedules these PRBs using a 1ms transmission time interval (TTI). The LTE MAC scheduler has a very complex way to assign resources to the associated users. Without digging into the details of the MAC scheduler operation, we focus on the last stage of resource assignment procedure in a certain TTI. On reaching that stage the MAC scheduler already builds up a list of users, which will be transmitting/receiving data in that TTI. For each user entry in the list, there is a corresponding value of the allocated number of PRBs, as well as the channel dependent Modulation and Coding Scheme (MCS) index. These two values are used to lookup the Transport Block Size (TBS) from a table defined in the 3GPP specifications [1]. This is a two dimensional table where each row representing one MCS index lists several values of TBS corresponding to the allocated number of PRBs. The obtained TBS value defines the size of the MAC frame transmitted to the user in that TTI. In this way, the user received throughput at the MAC layer in a certain TTI can be estimated if the TBS value for that user is known.

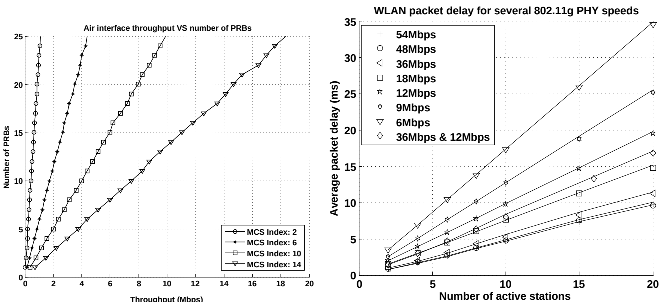
Fig. 1(a) shows that for a particular MCS index, the LTE throughput value has almost a linear relationship with the used number of PRBs. If described mathematically this relationship can be used to determine the required number of PRBs/TTI (q) to achieve a certain data rate R_i [kbit/sec] for a user having MCS index i. That is

$$q = \alpha_i \cdot R_i + \beta_i$$

α_i is slope of a straight line (as shown in 1(a)) described in units of PRBs/kbps. β_i is the intercept at the y-axis and has units of number of PRBs. It can be noticed that α_i is the data rate dependent part, while β_i is the data rate independent part of network resource requirement for a user with channel conditions mapped to MCS index i.

B. *Average user throughput estimation in WLAN network*

The user throughput in WLAN (IEEE 802.11) can be computed if the packet transmission delay is known. However, due to the random back-off time and possible packet collisions, the time required to transmit a packet successfully is highly variable. Moreover, this transmission time also depends on several factors like user channel conditions, number of users, user traffic demands etc. In [10], two dimensional Markov chain model has been used to compute the achievable throughput of 802.11b network with a certain number of stations having same channel conditions and traffic pattern. [11] has extended the model to calculate the average packet delay. The mathematical analysis in [11] assumes a network of n contending stations where each station has always a packet to transmit. The analysis yields two probability value: probability that there is at least one transmission in the considered



(a) (b)

Fig. 1. The left figure shows the relationship of LTE air interface user throughput and number of PRBs for different MSC index values. Each curve represents one MCS index. The right figure shows the packet delay estimation for users in WLAN network using proposed analytical model.

slot time (P_{tr}) and the probability that an occurring packet transmission is successful (P_s).

$$P_{tr} = 1 - (1-\tau)^n, \qquad P_s = \frac{n \cdot \tau \cdot (1-\tau)^{n-1}}{1-(1-\tau)^n}$$

where τ is the stationary probability that the station transmits a packet in a randomly chosen slot time. The relationships of P_{tr} and P_s are used to calculate the $E[slot]$ which is the average length of a slot time. The average length of a slot time is obtained considering that, with probability $1 - P_{tr}$, the slot time is empty; with probability $P_{tr}P_s$ it contains a successful transmission, and with probability $P_{tr}(1 - P_s)$ it contains a collision.

$$E[slot] = (1 - P_{tr})\sigma + P_{tr} \cdot P_s \cdot T_s + P_{tr}(1 - P_s) \cdot T_c \quad (1)$$

here σ is the duration of an empty slot time, T_s is the average time the channel is sensed busy because of a successful transmission, and T_c is the average time the channel is sensed busy by each station during a collision. Assuming $E[X]$ as the average number of slot times for a successful packet transmission, the value $E[X]$ can be found by multiplying the number of slot times the packet is delayed in each back-off stage by the probability to reach this back-off stage. The final form of $E[X]$ is as given below

$$E[X] = \frac{(1-2p)\cdot(W+1) + pW\cdot(1-(2p)^m)}{2\cdot(1-2p)\cdot(1-p)} \quad (2)$$

Finally, the average delay of a successfully transmitted packet $E[D]$ is given as following

$$E[D] = E[X] \cdot E[slot] \quad (3)$$

In equation (3), the values of σ, m and W can be obtained from the 802.11 specifications. The values of the other two unknown parameters i.e., T_s and T_c depend on fact whether the basic or RTS/CTS scheme is used. For example, with RTS/CTS scheme enabled these values are as follows

$$\begin{aligned} T_s^{rts} &= T_{RTS} + T_{SIFS} + \delta + T_{CTS} + T_{SIFS} + \delta \\ &+ T_H + T_{E[P]} + T_{SIFS} + \delta + T_{ACK} + T_{DIFS} + \delta \quad (4a) \end{aligned}$$

$$T_c^{rts} = T_{RTS} + T_{DIFS} + \delta \quad (4b)$$

where δ is the propagation delay and $T_H = T_{PHYhdr} + T_{MAChdr}$ is the time to transmit header data associated with PHY and MAC protocols and $T_{E[P]}$ is the time to transmit a data packet of mean size $E[P]$.

In the above described analysis, it has been assumed that all stations have same the channel conditions and therefore transmit with the same PHY data rate. In a realistic scenario this assumption cannot always be fulfilled. In order to use equation (3) in a scenario where users have different channel conditions and PHY data rates, it must be extended. The direct influence of the PHY data rate on average packet delay estimation can be observed in the computation of T_s and T_c (see equation (4)) where the user PHY data rate determines the value of $T_{E[P]}$, the time to transmit the data packet. Therefore, a network of users with different PHY data rate can be seen as a system with single server and multiple queues where the user PHY data rate is incorporated in the size of job. Excluding the medium contention time and assuming that a user always has a packet to transmit, the mean service time of such a system can be computed as

$$\hat{T}_{E[P]} = \frac{E[P]}{E[\text{PHY data rate}]}$$

Furthermore, when stations are transmitting at different PHY data rate, the transmission speed of control signals in the network, i.e., T_H, T_{RTS}, T_{CTS} and T_{ACK} is limited by the station having the lowest PHY data rate. This implies that the users must transmit control signals at a PHY data rate, which can be received by the all users. But, the data packet is transmitted by the user's own current PHY data rate. Incorporating the modified values of T_s and T_c in equation (1) produces $E[\widehat{slot}]$

$$E[\widehat{slot}] = (1 - P_{tr})\sigma + P_{tr}P_s\hat{T}_s + P_{tr}(1 - P_s)\hat{T}_c \quad (5)$$

As the value of $E[X]$ is independent of user PHY data rate, equation (3) takes the following form

$$\widehat{E[D]} = E[\widehat{slot}] \cdot E[X] \quad (6)$$

Figure 1(b) shows the average packet delay experienced by users transmitting in 802.11g network with RTS/CTS enabled. The solid lines show the estimated values using the equation (6). The markers on a solid line represents the delay values obtained from simulation results. It is evident from the figure that the modified model can precisely estimate the mean packet delay values for both scenarios, i.e., when all users have same PHY data rate as well as when users with different PHY data rate are mixed together. The mean packet delay value computed with (6) can be used to estimate the average user throughput Y in the network, i.e.,

$$Y = \frac{E[P]}{\widehat{E[D]}} = \frac{E[P]}{E[X] \cdot E[\widehat{slot}]} \quad (7)$$

IV. Optimized network resource allocation

When the network path costs for a multi-homed user are known, the problem of optimal resource utilization can be solved using mathematical techniques. In this work we have selected Integer Linear Programming (ILP) to solve this problem. A mathematical model for this purpose is discussed in first subsection. The second section explains how the non-linear relation for WLAN throughput computation is linearized and the third section shows how this model is integrated into simulation environment.

Given

U	a set of users
α_j	Data rate dependent part of the LTE link cost in PRBs per kbps for user j, for each $j \in U$
β_j	Data rate independent part of the LTE link cost in PRBs for user j, for each $j \in U$
ϕ_j	WLAN PHY data rate of a user j, for each $j \in U$
λ_j	Minimum data rate (kbps) demand of a traffic flow destined to user j, for each $j \in U$
Λ_j	Maximum data rate (kbps) allocation for a traffic flow destined to user j, for each $j \in U$
Ω	Number of available PRBs for the LTE access network
G	Mean packet size of active WLAN users in bit
$\widetilde{T_j}$	Per bit transmission delay excluding medium contentions for a user j with PHY data rate ϕ_j, for each $j \in U$

Defined variables

R_j	Size of sub-flow in kbps sent over the LTE access link to user j, for each $j \in U$
V_j	Size of sub-flow in kbps sent over WLAN access link to user j, for each $j \in U$
Y	Average throughput of active users in WLAN network in kbps
E_j	Auxiliary binary variable; its value for a user j is either 1 if $R_j > 0$ or 0 otherwise, for each $j \in U$
F_j	Auxiliary binary variable; its value for a user j is either 1 if $V_j > 0$ or 0 otherwise, for each $j \in U$

Maximize

$$\sum_{j \in U} R_j + Y$$

Subject to

1. $\sum_{j \in U} (\alpha_j \cdot R_j + \beta_j \cdot E_j) \le \Omega$
2. $\lambda_j \le R_j + V_j \le \Lambda_j, \quad \forall j \in U$
3. $\sum_{j \in U} F_j \ge 1$
4. $\sum_{j \in U} F_j \cdot V_j = \sum_{j \in U} F_j \cdot Y$
5. $V_j \le G/\widetilde{T_j}, \quad \forall j \in U$
6. $0 \le R_j \le \Lambda_j, \quad \forall j \in U$
7. $0 \le Y_j \le \Lambda_j, \quad \forall j \in U$

Fig. 2. Mathematical model for the resource allocation in algebraic form

A. Mathematical model for resource allocation

Fig. 2 shows the formulation of the problem in algebraic form. The model defines U as the set of multi-homed users. Each element of this set has a number of input parameters, e.g., network path cost for LTE (α,β) and user WLAN PHY data rate (ϕ) according to the user channel conditions in the corresponding network. The maximum and minimum range of user data rate demands (λ, Λ), which is based on the individual user application. The amount of available network resources in LTE (Ω) are also considered as input parameters. The output parameters for each user in set U include the assigned data rate over the LTE network and the WLAN network paths (R,V). It is obvious that the goal of this model is to achieve the highest possible spectral efficiency from the two network access technologies. The higher the spectral efficiency, the higher the network throughput. Hence the objective is to maximize the user data rate over the two network paths, i.e., R and V for every multi-homed user. The model imposes several constraints, however for the sake brevity the most important seven constraints are listed in Fig.2. The first constraint ensure that the available LTE network resources should not be exceeded when allocating the data rates for users. The second constraint dictates that the user data rate allocation should lie within the specified range. The third constrains enforce the use of WLAN network by at least one user. The 4th constraint allows the users to distribute the available WLAN network throughput according to their needs. This means if a user does not always have some data to transmit his throughput share can be used by other users. Nevertheless, a user cannot transmit at higher data rates than allowed by his PHY data rate as seen in constraint 5. According to the requirements of goal of optimized resource allocation a user may receive multiple sub-flows or one single flow of application data as shown in constraint number 6 and 7.

B. Linearizing WLAN throughput estimation formula

The mathematical model has to rely on equation (7) for the estimation of the average user throughput in WLAN network. However, it is clearly a nonlinear relation, which must be linearized using some work around to use in linear programming. For this purpose, the equation (5) is split into two parts. One part depends only on n, which represents the total number of active stations in WLAN network. The other part incorporates both n as well as $\hat{T}_s$ variables. As RTS/CTS is enabled for all users therefore $\hat{T}_c = T_c$.

$$\widehat{E[slot]} = f_x(n) + f_y(n) \cdot \hat{T}_s$$

where $\quad f_x(n) = (1-P_{tr})\sigma + P_{tr}(1-P_s) \cdot T_c, \quad f_y = P_{tr} \cdot P_s$

Moreover, equation (2) shows that $E[X]$ is a function of only one variable n, which allows us to write

$$\widehat{E[D]} = E[X] \cdot \widehat{E[slot]} = f_1(n) + f_2(n) \cdot \hat{T}_s \qquad (8)$$

where $\quad f_1(n) = E[X] \cdot f_x(n),$ and $\quad f_2(n) = E[X] \cdot f_y(n)$

In order to simplify the relation in equation (8), $f_1(n)$ and $f_2(n)$ are approximated using 3rd order polynomial curve fitting as shown below.

$$f_1(n) \approx A_{11}n^3 + A_{12}n^2 + A_{13}n + A_{14},$$

$$f_2(n) \approx A_{21}n^3 + A_{22}n^2 + A_{23}n + A_{24}$$

where all occurrences of A represent constant value numbers. Fig. 3 shows that the curve fitting process generates an accurate enough approximation for $f_1(n)$ & $f_2(n)$ functions with norm of residuals as 3.9×10^{-5} and 0.14 respectively. Using the approximate function of $f_1(n)$ and $f_2(n)$ in equation

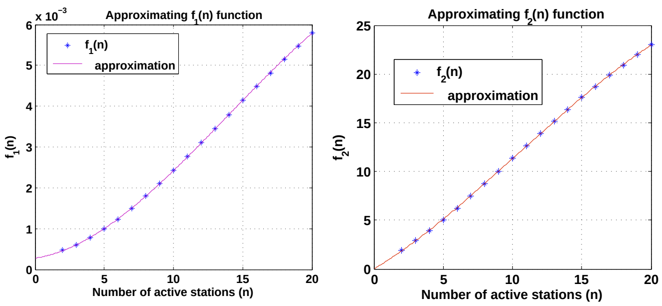


Fig. 3. Approximation of $f_1(n)$ and $f_2(n)$ using polynomial curve fitting

(8) & (7) and after a few manipulation steps of algebra we get

$$n \cdot G = Y \cdot \{A_{11}n^4 + A_{12}n^3 + A_{13}n^2 + A_{14}n + [A_{21}n^3 + A_{22}n^2 + A_{23}n + A_{24}] \cdot \sum_{i=1}^{n} T_{s_i}\} \quad (9)$$

The variable n in above equation can be replaced with a summation of binary variables F_i, which represents whether a station i is active or not. If there are total Z number of users in WLAN network out of which only n users are active then $n = \sum_{i=1}^{Z} F_i$. Similarly higher order variables of n can linearized as following

$$n^2 = (\sum_{i=1}^{Z} F_i)^2 = \sum_{i,j=1}^{Z} F_i \cdot F_j = \sum_{i,j=1}^{Z} \chi_{i,j}^{\text{F2}}$$

Where $\chi_{i,j}^{\text{F2}}$ represents a product of two binary variables F_i & F_j. The value of $\chi_{i,j}^{\text{F2}}$ is determined by following three constraints.

$$\chi_{i,j}^{\text{F2}} \leq F_i, \quad \chi_{i,j}^{\text{F2}} \leq F_j, \quad \chi_{i,j}^{\text{F2}} \geq F_i + F_j - 1$$

Moreover, the continuous variable Y can be multiplied with binary product variable $\chi_{i,j}^{\text{F2}}$ to get the product term $\chi_{i,j}^{\text{YF2}}$, i.e.,

$$Y \cdot n^2 = Y \cdot \sum_{i,j=1}^{Z} \chi_{i,j}^{\text{F2}} = \sum_{i,j=1}^{Z} \chi_{i,j}^{\text{YF2}}$$

Taking $\breve{Y}$ as the maximum value of Y the following three constraints help determine the value of product term $\chi_{i,j}^{\text{YF2}}$

$$\chi_{i,j}^{\text{YF2}} \leq \breve{Y} \cdot \chi_{i,j}^{\text{F2}}, \quad \chi_{i,j}^{\text{YF2}} \leq Y, \quad \chi_{i,j}^{\text{YF2}} \geq Y - \breve{Y} \cdot (1 - \chi_{i,j}^{\text{F2}})$$

The summation term $\sum T_{s_i}$ in equation (9), which represents the addition of T_s from all active users, can be written as following

$$\sum_{i=1}^{n} T_{s_i} = \sum_{i=1}^{Z} F_i \cdot T_{s_i}$$

Adopting this strategy equation (9) can be linearized as following

$$\sum_{i=1}^{n} F_i \cdot G = \sum_{j,k,l,m=1}^{Z} (A_{11} + A_{21} \cdot T_j) \cdot \chi_{j,k,l,m}^{\text{YF4}} + \sum_{j,k,l=1}^{Z} (A_{12} + A_{22} \cdot T_j) \cdot \chi_{j,k,l}^{\text{YF3}} + \sum_{j,k=1}^{Z} (A_{13} + A_{23} \cdot T_j) \cdot \chi_{j,k}^{\text{YF2}} + \sum_{j=1}^{Z} (A_{14} + A_{24} \cdot T_j) \cdot \chi_j^{\text{YF}} \quad (10)$$

It should be noted that equation (10) is valid for $n > 1$. If there is only one active user in the system then no medium contention would take place. In that particular case

$$E[D] = T_s + T_{\text{back-off}} = T_s + \frac{W-1}{2} \cdot \sigma = \widetilde{T} \quad (11)$$

Equation (10) and equation (11) can be combined by introducing another binary variable L, which is 1 if there is only one active user and 0 otherwise. The value of L is determined by following constraints

$$2 - L \cdot 10^9 \leq \sum_{j=1}^{Z} F_j \quad \text{and} \quad 1 + (1 - L) \cdot 10^9 \geq \sum_{j=1}^{Z} F_j$$

Finally, the linearized version of equation (9), which is valid for $n \geq 1$ is given as below

$$\sum_{i=1}^{n} F_i \cdot G = \sum_{j,k,l,m=1}^{Z} (A_{11} + A_{21} \cdot T_j) \cdot \chi_{j,k,l,m}^{\text{YF4}} + \sum_{j,k,l=1}^{Z} (A_{12} + A_{22} \cdot T_j) \cdot \chi_{j,k,l}^{\text{YF3}} + \sum_{j,k=1}^{Z} (A_{13} + A_{23} \cdot T_j) \cdot \chi_{j,k}^{\text{YF2}} + \sum_{j=1}^{Z} (A_{14} + A_{24} \cdot T_j) \cdot \chi_j^{\text{YP}} - \sum_{j=1}^{Z} \chi_j^{\text{YFL}} \cdot (A_{11} + A_{12} + A_{13} + A_{14} + (A_{21} + A_{22} + A_{23} + A_{24}) \cdot T_{s_j} - \widetilde{T_j}) \quad (12)$$

C. Applying the mathematical model in a simulation scenario

In the investigated scenario the LTE coverage is available in the whole area of user movement while WLAN coverage is limited in a circular area of 100 meter radius around a hotspot. This implies that the users always have LTE access available and WLAN coverage is only found in the vicinity of the hotspot (see Fig.4). During the resource assignment process, the flow management function classifies users into the following three categories (i) users with LTE access only and running VoIP or video applications (ii) users with LTE and WLAN access running any type of application (iii) users running FTP or HTTP applications with LTE access only. Users in the first category must be assigned the required minimum data rate through LTE as there is no other access available for them. Users in the second category are multi-homed users whose data rate will be decided by the aforementioned mathematical model. For users belonging to the third category, they must

 ISBN: 978-1-61208-229-5

get their traffic through the LTE path, however, it is not clear how much data rate should be allocated to them in order to achieve the optimized resource allocation objective. This issue is resolved by using the following work around: the users are assigned a $\widetilde{T_j}$ value greater than unity and they are put into the second category. The value of $\widetilde{T_j}$ greater than unity will refrain the LP solver to assign any data rate for these users over the WLAN path while the data rate for the LTE path will be decided based on the global objective of the optimized resource allocation. It is assumed here that each

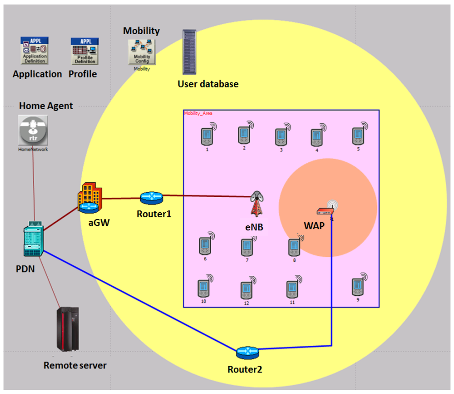


Fig. 4. Simulation scenario overview in the OPNET simulator. The large and small circular areas show the coverage of LTE and WLAN networks respectively. The user movement is restricted to the rectangular area.

user is running only one application. For a constant bit rate application, e.g., VoIP or video the minimum data rate is set equal to the maximum data rate in the model input parameters. For TCP based flows these two values can be set according to the network operator's policy. It should be noted that the problem has been formulated in a way that it guarantees the minimum data rate for all users and then assigns an additional data rate up to the maximum data rate while optimizing the spectral efficiency of the access networks.

The resource assignment process by the flow management function is carried out periodically in order to adapt to any changes in the user channel conditions. For this purpose, user channel condition parameters are obtained through cross layer information from the base stations of the two access technologies. According to this updated information, the mathematical model may reevaluate the solution considering the updated input parameters. As described earlier, the mathematical model is formulated using linear programming and solved using the C application programming interface (API) of ILOG CPLEX from IBM [7] which has been integrated inside the OPNET simulator by the authors. The output of this process consists of user data rates on each network path. These decided data rates are then conveyed to the users using fast LTE control plane signalling.

V. Simulation Results

This section shows the benefits of the proposed approach with the help of simulation results. For this purpose two scenarios are considered. In one scenario users do not make

TABLE I
Simulation configurations

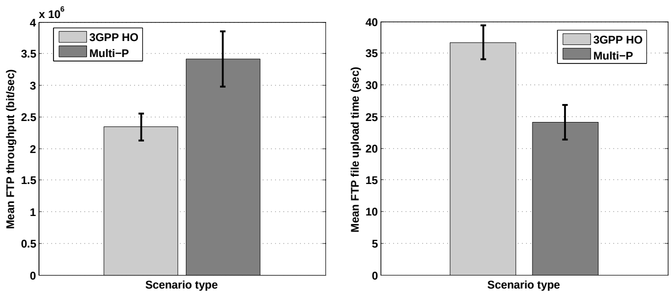
Parameter	Configurations
Total Number of PRBs	25 PRBs (5 MHz specturm)
Mobility model	Random Direction (RD) with 6 km/h
Number of users	3 VoIP, 2 Skype video call, 7 FTP uplink users
LTE Channel model	Macroscopic pathloss model , Correlated Slow Fading [13]
LTE MAC Scheduler	TDS: Optimized Service Aware, FDS: Iterative RR approach [12]
WLAN technology	802.11g, RTS-CTS enabled, coverage $\approx$ 100 m
VoIP traffic model	G.722.2 wideband codec, 23.05kbps data rate
Skype video model	MPEG-4 codec, 512kbps, 640x480 resolution, 30fps, play-out delay: 250 ms
FTP traffic model	FTP File size: constant 10 MByte continuous file uploads one after the other.
Simulation run time	10^3 seconds,10 seeds, 95% confidence interval

simultaneous use of LTE and WLAN access technologies. Instead the user traffic is completely handed over to WLAN as soon as the user is in the hotspot coverage, otherwise all traffic takes its path through the LTE access. This is the default policy for a multi-homed user according to the 3GPP specifications and therefore it will be referred to as "3GPP HO" case. Whereas, the second scenario extends the 3GPP architecture to support the simultaneous use of wireless interfaces, this will be referred to as "Multi-P". In this case, user traffic flows are distributed over the WLAN and the LTE access network. The traffic flow distribution policy is derived from the output of the optimization problem solved using linear programming. As a result, a user traffic flow is either sent over one network path with the least cost or it is split into two appropriately sized sub-flows each taking one network path to the destination.

Fig. 4 shows an overview of the simulation network model implemented in OPNET. The system is populated with 12 users generating a rich traffic mixture of: Voice over IP (VoIP), uplink File Transfer Protocol (FTP), and video conference (i.e., Skype video call). The users move within one LTE eNB cell, and within this cell one wireless access point (or hotspot) is present. The simulation configuration parameters are shown in Table I. Besides, in the "Multi-P" scenario the minimum data rate for FTP users is assigned as 200kbps while the maximum data rate limit is set to a very high value of 25Mbps.

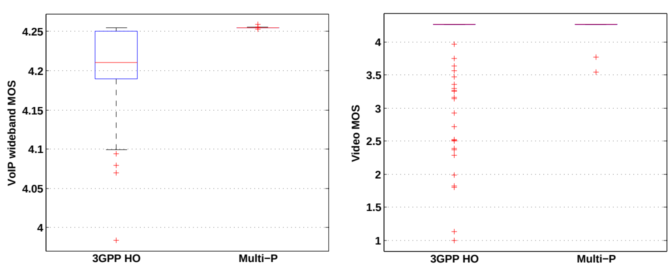
In "3GPP HO" scenario, the users make handover between two access technologies without following make-before-break approach, i.e. the connection is broken from one network, and a new connection is established to the other one. Though MIPv6 keeps all IP layer connections alive through seamless handover, the user might lose data packets buffered at the lower protocol layers of the previously in use network interface. For example, LTE buffers the received IP packets at PDCP, RLC and MAC layers while WLAN keeps all the data buffered at MAC layer before transmission over the radio interface. Therefore, when making complete handover from one access technology to another, this buffered data is discarded and have to be recovered by upper layers through retransmissions. This behavior leads to applications performance degradation for both TCP and UDP based applications.

The "Multi-P" scenario the users are allowed to use the WLAN access when it is in the coverage, and can still keep the LTE connection alive and use it at the same time. In the coverage of WLAN access, the flow management client function sends user traffic on WLAN link only when user PHY mode is 9Mbps or higher. This is because when a user enters in 6Mbps mode it implies that the user is almost at edge of coverage which is a strong indication that loss of WLAN link is imminent. Hence, no new traffic data is scheduled for WLAN link which gives user a chance to transmit already buffered data to the access point before the loss of link happens. Moreover, "Multi-P" approaches, in contrast to "3GPP HO" scenario, keep buffered data at the minimal required level through the use of network path capacity estimations.



(a) Mean user FTP uplink throughput (b) Mean user file upload time

Fig. 5. User evaluation for non-real time application



(a) VoIP call quality (b) Video call quality

Fig. 6. User evaluation for real time applications

Fig. 5 shows the performance of FTP uplink application for the users. It can be seen that the "Multi-P" algorithm achieves the best results for FTP uplink user traffic. The figure shows that "Multi-P" provides 32% higher user FTP throughput than "3GPP HO" scenario. The higher throughput helps users finish file upload faster. The gain in FTP throughput is mainly coming from the proper management of network resources where users with good channel conditions are assigned more resources in LTE network. Similarly in WLAN network users with the best channel condition are allowed to transmit.

Fig. 6 shows the boxplot of user Mean Opinion Score (MOS) values of VoIP and video services. The MOS values of the wideband VoIP codec and video codec are computed using the modified E-model and Evalvid toolkit as described in [4] and [3], respectively. The figure shows that VoIP and Video users in "Multi-P" scenario mostly achieve the best possible MOS value. The users in "3GPP HO" scenario often suffer from quality loss when making handover as well as when transmitting over QoS unaware WLAN network. This shows that proper management of network resources as performed by flow management function not only improves the network capacity but also enhances user QoE.

VI. Conclusion

This work highlighted the importance of multi-homing support in integrated heterogeneous wireless networks of 3GPP and non-3GPP access technologies. The existing 3GPP specifications for the integration of two types of the access technologies (i.e. 4G LTE and WLAN) are extended following IETF standards to realize multi-homing support for the users. This work mainly focuses on the problem of optimum resource utilization in such a heterogeneous network where the users and network operators can take advantage of multi-homing support. The problem of optimum network resource allocation is mathematically modeled using the linear programming technique. The mathematical model is then integrated in the network simulator to decide the network resource allocation for multi-homed users during the simulation. With help of simulation results it is shown that the proposed scheme of resource allocation brings twofold gain when compared to the 3GPP proposal. On the one hand it significantly improves the network capacity and on the other hand it fulfills the user application QoS demands which otherwise cannot be satisfied from QoS unaware non-3GPP access technologies.

References

[1] 3GPP Technical Report TS 36.213, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures", v10.2.0, Jun. 2011.
[2] Toseef, U., Zaki, Y., Timm-Giel, A., and Görg, C., "Development of Simulation Environment for Multi-homed Devices in Integrated 3GPP and non-3GPP Networks", The 10th ACM International Symposium on Mobility Management and Wireless Access in Paphos, Cyprus, Oct. 2012.
[3] J.Klaue, B. Rathke, and A. Wolisz, "EvalVid - A Framework for Video Transmission and Quality Evaluation", In Proc. of the 13th International Conference on Modeling Techniques and Tools for Computer Performance Evaluation, pp. 255-272, Illinois, USA, Sept. 2003.
[4] Toseef, U., Li, M.,Balazs, A., Li,X., Timm-Giel A., and Görg, C., "Investigating the Impacts of IP Transport Impairments on VoIP service in LTE Networks," in 16. VDE/ITG Fachtagung Mobilkommunikation, Osnabrück, Germany, May 18-19, 2011.
[5] 3GPP Technical Report TS 23.402, "Architecture enhancements for non-3GPP accesses, 3rd Generation Partnership Project", v10.6.0, Dec. 2011.
[6] OPNET, http://www.opnet.com, as accessed in Oct. 2012.
[7] CPLEX, http://www.ibm.com/software/, as accessed in Oct. 2012.
[8] Yasir Zaki, "Long Term Evolution (LTE) Model Development Within OPNET Simulation Environment", OPNET Work Dec. 2011, Washington.
[9] G. Bianchi, "Performance Analysis of the IEEE 802.11 Distributed Coordination Function," IEEE Journal on Selected Areas in Communications, Vol. 18, No. 3, pp. 535-547, Mar. 2000.
[10] R. Litjens, F. Roijers, J. L. van den Berg, R. J. Boucherie, and M. Fleuren, "Performance Analysis of wireless LANs: an Integrated Packet/Flow Level Approach", ITC Conference, Berlin, Aug. 2003.
[11] P. Chatzimisios, V. Vitsas, and A.C. Boucouvalas, "Throughput and Delay Analysis of IEEE 802.11 Protocol", in Proceedings of the 5th IEEE International Workshop on Networked Appliances (IWNA 2002), pp. 168-174, Liverpool, UK, 30-31 Oct. 2002.
[12] S. N. K. Marwat, "Performance Evaluation of Bandwidth and QoS Aware LTE Uplink Scheduler",10th International Conference on Wired/Wireless Internet Communications, Santorini, Greece, Jun. 2012.
[13] 3GPP Technical Report TS 25.814, "Physical layer aspects for E-UTRA", 3rd Generation Partnership Project, v.7.1.0, Sept. 2006.

 ISBN: 978-1-61208-229-5