SHARING

SELF-ORGANIZED HETEROGENEOUS ADVANCED RADIO NETWORKS GENERATION

Deliverable D4.3

Inter-system offloading: innovative concepts and performance evaluation

Date of delivery	25/12/2014
Contractual date of delivery	31/12/2014
Project number	C2012/1-8
Editor(s)	Engin Zeydan (AVEA)
Author(s)	Engin Zeydan (AVEA), Ahmet Serdar Tan (AVEA), Mehdi Bennis (UOULU), Javier Valiño (TTI)
Dissemination level	PU
Workpackage	4
Version	V1.0
Total number of pages	41

Abstract:

This deliverable provides innovative concepts and solutions regarding intelligent inter-system offloading within heterogeneous network. The deliverable provides details and performance evaluations of developed solutions for inter-system offloading.

The solutions are presented for heterogeneous networks within the topics of: Inter-LTE traffic offloading via middleware deployment, seamless offloading, capacity aware multi-user offloading, joint offloading and scheduling for dual mode small cells.

Keywords:

Cellular Networks, Heterogeneous Networks, LTE, WiFi, Small Cells, Offloading, Load Balancing

Document Revision History

Version	Date	Author	Summary of main changes
0.1	30/09/2014	Engin Zeydan	Creation of the Table of Contents (ToC)
0.2	27.10.2014	Javier Valiño (TTI)	Added Inter-LTE traffic offloading via middleware deployment section
0.3	27.10.2014	Engin Zeydan (AVEA), Mehdi Bennis (UOULU)	Added Seamless offloading in heterogeneous wireless networks, Capacity Aware Multi-User Load Balancing for Heterogeneous Wireless Networks sections, Joint offloading and scheduling strategies for dual mode small cells sections
0.4	12.11.2014	Engin Zeydan (AVEA)	Added Executive Summary, Introduction and Conclusion
0.5	01.12.2014	Engin Zeydan (AVEA)	Updated based on reviewer comments (TTI,FT)
0.6	19.12.2014	Engin Zeydan (AVEA)	Updated based on TM and PC
1.0	25.12.2014	Engin Zeydan (AVEA)	Final version

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This deliverable provides solutions of SHARING in Work Package 4 Task 4.2 for inter-system radio access offloading and related performance evaluations. The objective of Work Package 4 is to identify new opportunities and challenges offered by small cells (pico-cells and femto-cells) and WiFi. Another objective of Work Package 4 is to propose and compare deployment strategies, load balancing algorithms and innovative interference management techniques in a way that optimizes energy efficiency and accounts for practical issues such as the limited capacity for signalling between nodes. Additionally, Work Package 4 conducts the pre-standardization research for convergence of Long-Term-Evolution (LTE) with other Radio Access Technologies (RATs) and investigates mechanisms allowing interoperability, as well as their ability to increase the capacity of the network and to offload part of the traffic.

The topic of this deliverable, inter-system offloading, which is an efficient and cost-effective integration of cellular and WiFi technologies, has recently attracted significant interest from academia, industry, and standardization bodies alike. On one hand, the inherent constraints of small cell networks, particularly due to cross-tier and co-tier interference, motivate offloading some of the traffic to the WiFi band to alleviate interference and ease congestion. On the other hand, due to the uncontrolled and unlicensed nature of WiFi, the competition for resources among a large number of hotspot users can yield to dramatically poor throughput. In such a scenario, offloading some of this traffic to a well-managed small cell network operating over the licensed spectrum can improve the performance.

In SHARING Work Package 4 Task 4.2, first, intelligent user allocation in heterogeneous networks via a middleware operating at IP level able to share context aware information regarding mobile users is studied. The results yield to a consistent QoS enhancement, indicating the advantages of the middleware deployment and the sharing of context user data. Task 4.2 also studies viable seamless offloading solutions to switch between 3GPP data networks and WiFi networks by embedding a QoE metric for handover decision into a multiple attribute decision making (MADM) algorithm. The utilized MADM algorithm together with a standards based handover mechanism is shown to initiate seamless handovers for video streaming scenarios. Another study in Task 4.2 is on load balancing between 3GPP (LTE) and WLAN networks using a novel capacity-aware multi-user MADM algorithm. It is shown that by using the capacity-aware multi-user MADM algorithm total channel utilization of the heterogeneous access network increases compared to single-user decision algorithms. The final study in Task 4.2 is on a cross-system learning approach for joint scheduling of the wireless resources between multi-mode small cell base stations, SCBSs, which are capable of transmitting simultaneously on both licensed and unlicensed bands. The results of the study show that the proposed cross-system learning approach converges to the optimal solution in less iterations compared to state-of-the-art approaches.

This deliverable builds on top of the concept descriptions and initial evaluation results provided in SHARING deliverable D4.1 [ShD41] and provides details of the abovementioned offloading research activities targeting resource optimization in heterogeneous networks. The results presented in this deliverable will be exploited in D4.6 in which the most promising concepts and techniques of Work Package 4 will be further investigated.

1 INTRODUCTION

The traffic volume and number of subscriptions in mobile networks are expected to rise rapidly due to the evolution of mobile terminals and the increased number of heavy-traffic services [Cis14], [Eri13]. The M2M and cloud-based services are expected to accelerate the subscriber and traffic growth even further.

In order to cope with the traffic deman of mobile users, mobile operators need to come up with alternative and efficient solutions. Network densification is one of the ways to increase the capacity of a mobile network. The traditional way to densify mobile networks has been to deploy new macro cells, either by adding new sectors to existing sites, or by deploying new macro sites. The benefit of a densified macro deployment is that the network performance can be improved with a relatively small amount of required new hardware, or new sites. However, as new macro sites are becoming increasingly difficult and often expensive to deploy, at least within urban environments, focus is put on the efforts to find more cost-efficient ways to densify the current networks.

An alternative to deploying new macro sites is to deploy low-power sites within traffic hotspots, i.e. the introduction of heterogeneous network deployments. In case of the heterogeneous network deployment, macro cells will provide wide area coverage, while the small low-power cells deployed within traffic hotspots will take care of the majority of the traffic volume. The downside of heterogeneous network deployments compared to the densified macro deployments is that a considerably larger number of new cells are required to be able to offer the same system performance. Even though the cost of a low-power site will typically be lower than the cost of a macro site, the overall situation may turn out to be quite challenging from the total network cost point of view.

SHARING D4.2 deliverable will look into many of the topics related to inter-RAT traffic offloading such as traffic offloading between LTE and WiFi and will evaluate the impacts of inter-RAT traffic offloading. It will investigate new ways to perform inter-RAT traffic offloading as efficiently as possible, with topics including concepts related to traffic offloading management via middleware deployment techniques, seamless load balancing techniques in multi-user scenarios, joint scheduling and offloading strategies. This deliverable discusses performance and deployment strategies of heterogeneous network deployments within different scenarios. The evaluation results presented in this deliverable indicate that the proposed inter-system traffic offloading mechanisms are able to improve both the user quality and the overall system capacity.

SHARING D4.2 deliverable will feed deliverable D4.6 which will summarize the work done within Work Package 4 during the project and describe the most promising concepts and techniques from tasks T4.1-T4.4 taking into account the results in D4.2-D4.5, as well as the results obtained after the mid-term summaries.

This deliverable presents an initial view on new opportunities, challenges and innovative concept candidates for heterogeneous network deployments. Chapter 2 presents Inter-LTE traffic offloading via middleware deployment techniques, Chapters 3-4 present concepts related to Seamless offloading techniques including capacity aware multi-user load balancing techniques in heterogeneous wireless networks. Decentralized and dynamic traffic offloading framework, in which small cells seamlessly steer their traffic between cellular and WiFi RATs is proposed. Chapter 5 discusses joint offloading and scheduling strategies for dual mode small cells.

2 INTER-LTE TRAFFIC OFFLOADING VIA MIDDLEWARE DEPLOYMENT

The goal is to study and describe technically a middleware operating at IP level on the control plane able to share context aware information regarding mobile users (not available nowadays) with the network, enabling more intelligente user allocation in cells and taking into account both 3GPP and WiFi APs. The performance of the middleware will be evaluated using system level simulations in order to assess the theoretical gains.

2.1 Solution Description

SHARING D4.2 deliverable (section 3.2) introduces the middleware at control plane and IP level needed to exchange user context information so as to intelligently allocate users on the most appropriate cell within the coverage area. This method makes use of the same idea, changing the scenario covered by D4.2 (just LTE macro cell deployments) with new WiFi small cells on top of the macro scenario.

The middleware basis relies on the idea that it is possible to introduce new information regarding user characteristics so as to help perform allocations. With this information and the data know by the network in terms of BS deployments at each point, it is possible to enhance the way users are attached to the network. For the sake of simplicity, please refer to D4.2 for further details about the middleware definition and strategies and assumptions considered. This chapter will just introduce the results considering the new scenario.

2.2 Scenario

The general layout for this simulation considers a grid of 5x5 macro BS able to handle 250 users and 120 Mbps each. The BS coverage radius is 50 meters and the Inter-Site Distance (ISD) is thus established at $\sqrt{3}\cdot 50$ meters. Extra WiFi access points are also considered, up to 50 users and 500 Mbps with 20m coverage radius.

The same conservative and realistic scenario assumptions were applied for running the simulations, obtaining the following layout coverage figures (Figure 12, Figure 2, Figure 3 and Figure 4).

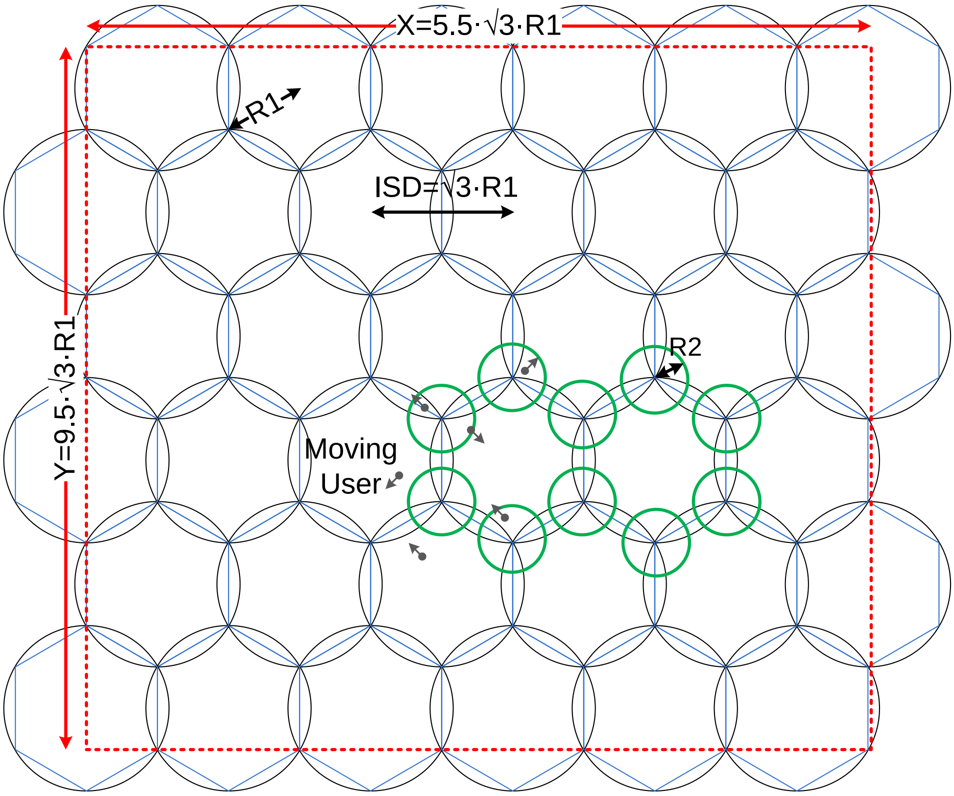


Figure 1 Conservative layout.

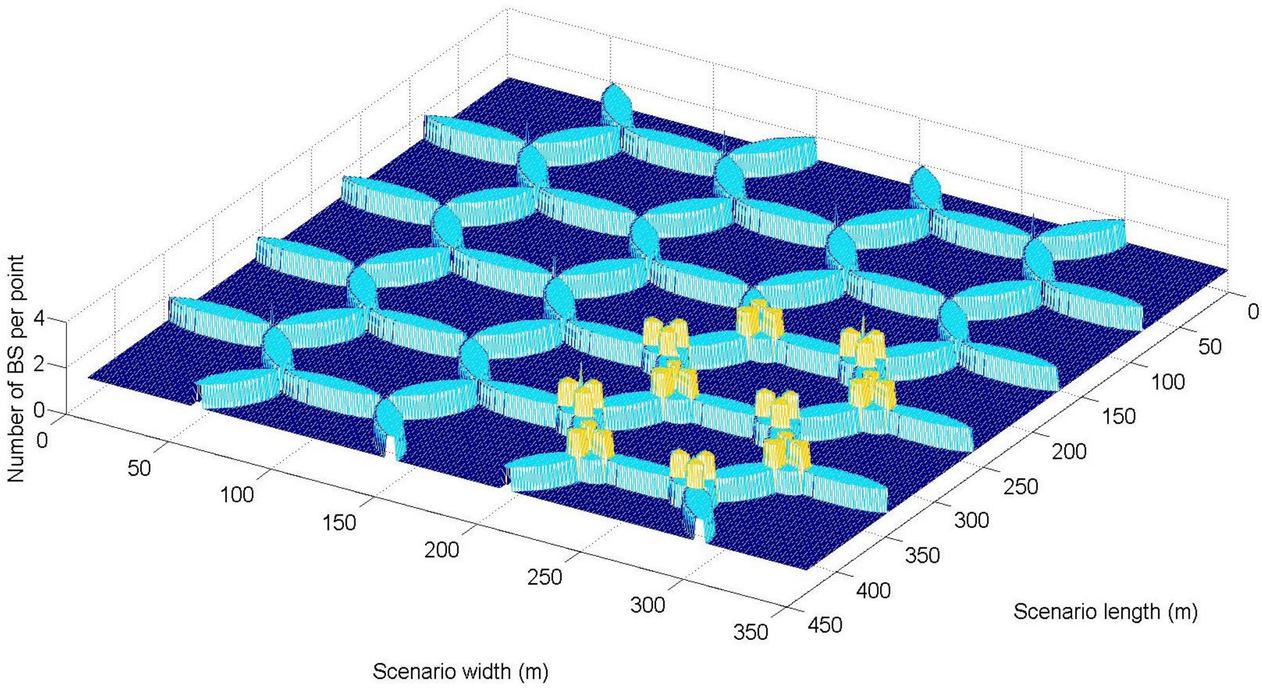


Figure 2 Number of BS seen at each point with conservative approach.

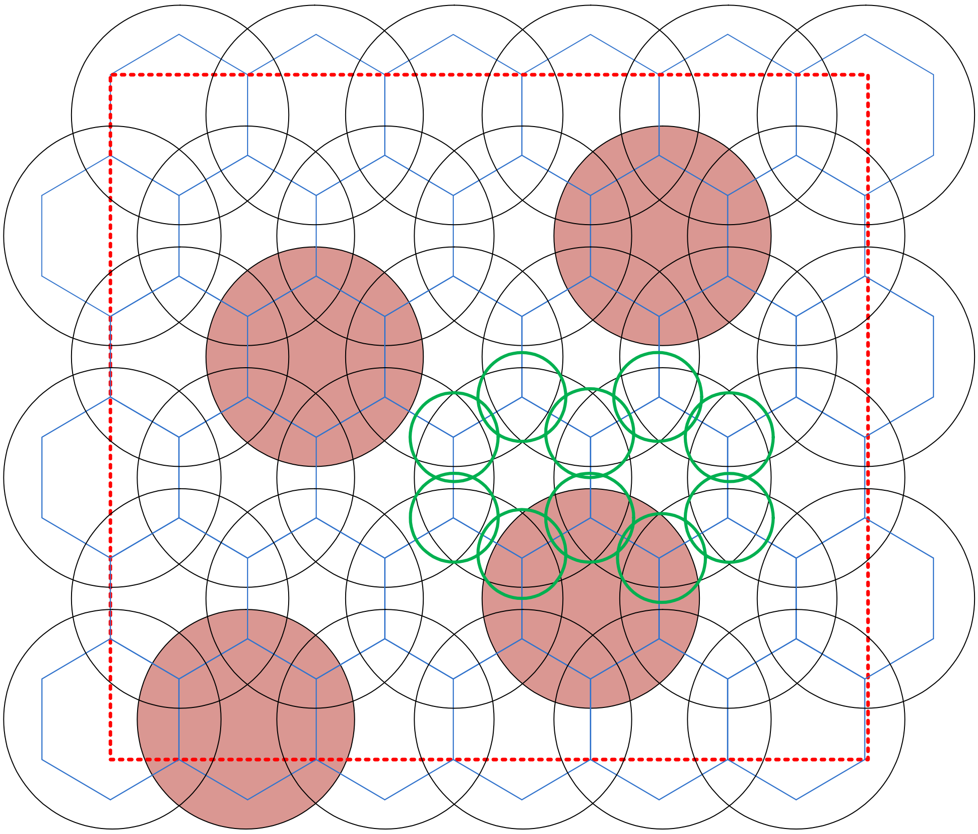


Figure 3 Realistic approach (1.5 coverage radius extended).

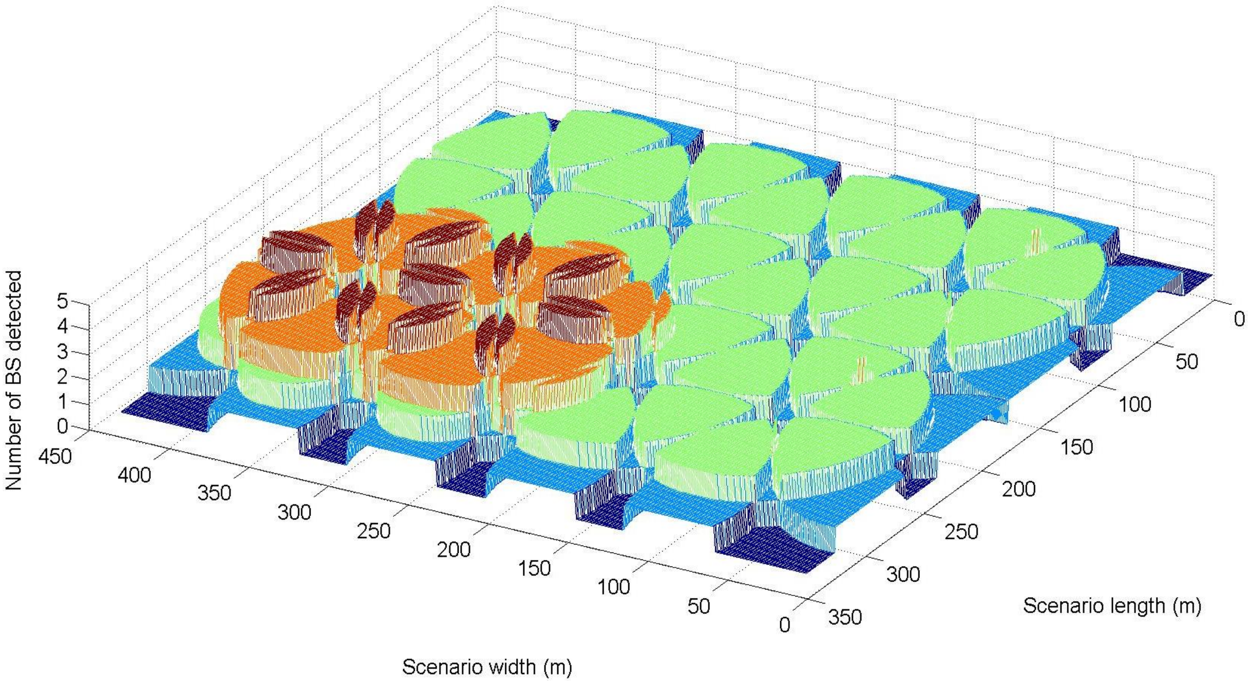


Figure 4 Number of BS seen at each point with realistic approach.

As a hot-zone is defined via the deployment of small cells, users have 20% more probability to be placed at that region.

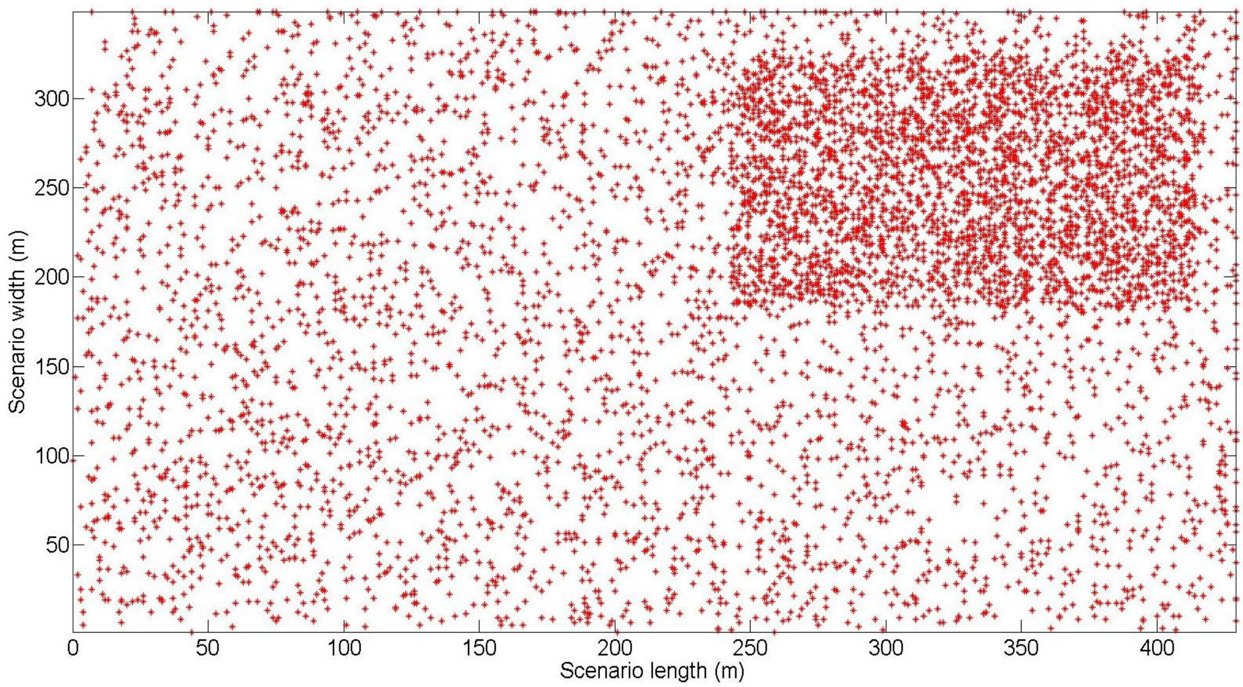


Figure 5 User position at peak hour.

2.3 Results

Depending on the policy and layout used, the following results are obtained:

- Any SHARING defined policies, intermediate or complex, obtains better QoS performance with respect to the simple approach (not using the new available data).
- The conservative scenario always produces lower QoS performance compared with the reaslstic scenario, as it allows less flexibility in terms of cell load balancing.
- The complex policy enables maintaining 100% QoS levels almost always. Just the peak hour, as it was dimensioned to overcome the network capacity, presents peaks of QoS below 100%.
- Using complex policy combined with conservative scenarios, results in 15% peak and 30% mean enhancements in terms of overall QoS with respect to simple policy.

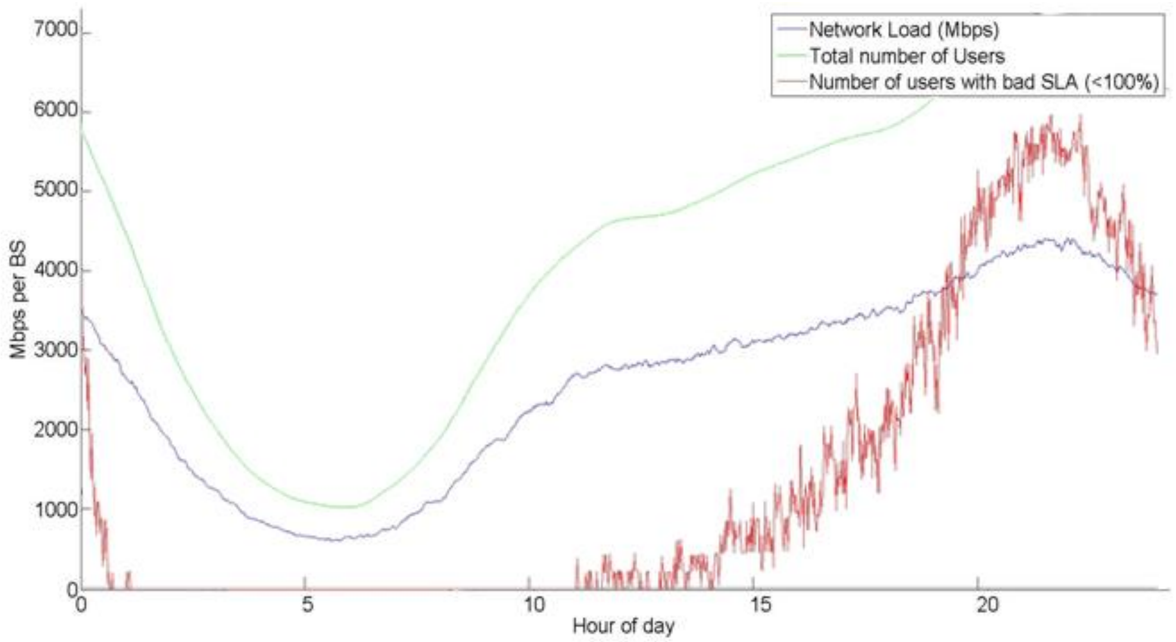


Figure 6 QoS (red curve) of the network using simple policy in conservative scenario.

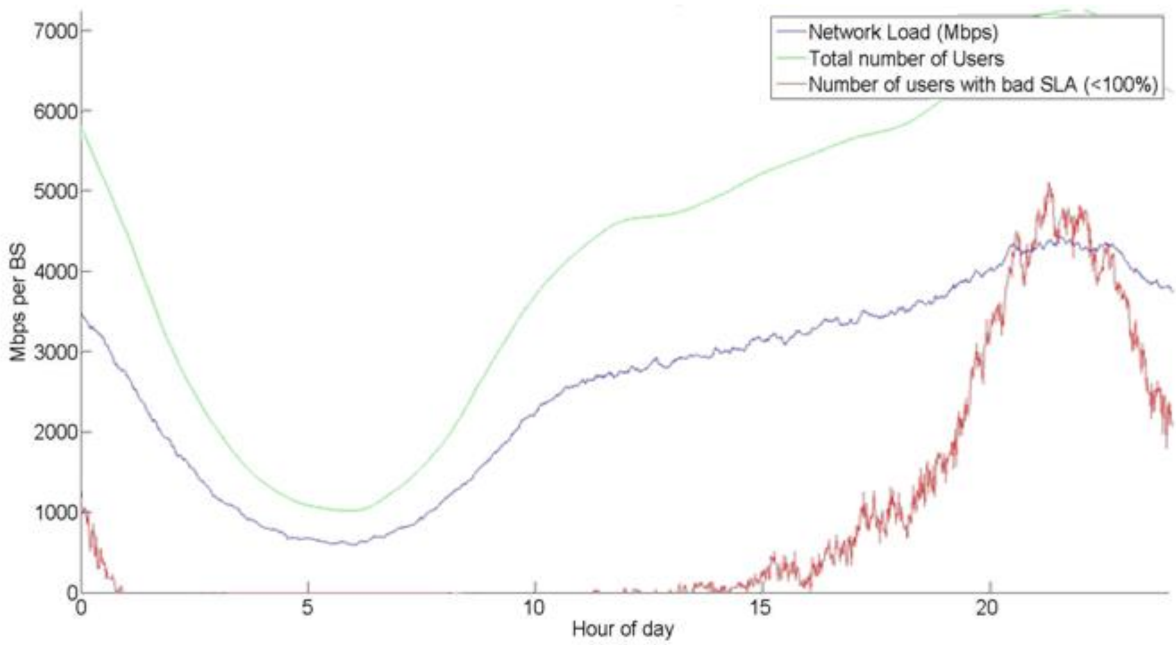


Figure 7 QoS (red curve) of the network using complex policy in conservative scenario.

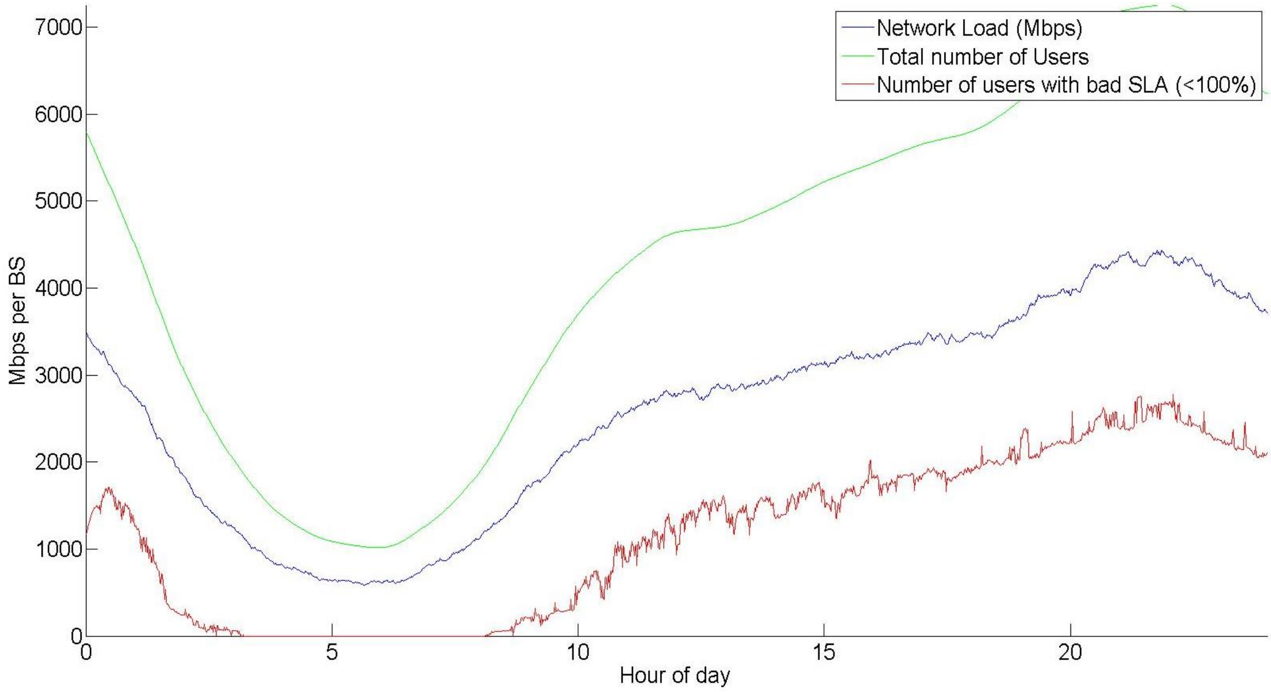


Figure 8 QoS (red curve) of the network using simple policy in realistic scenario.

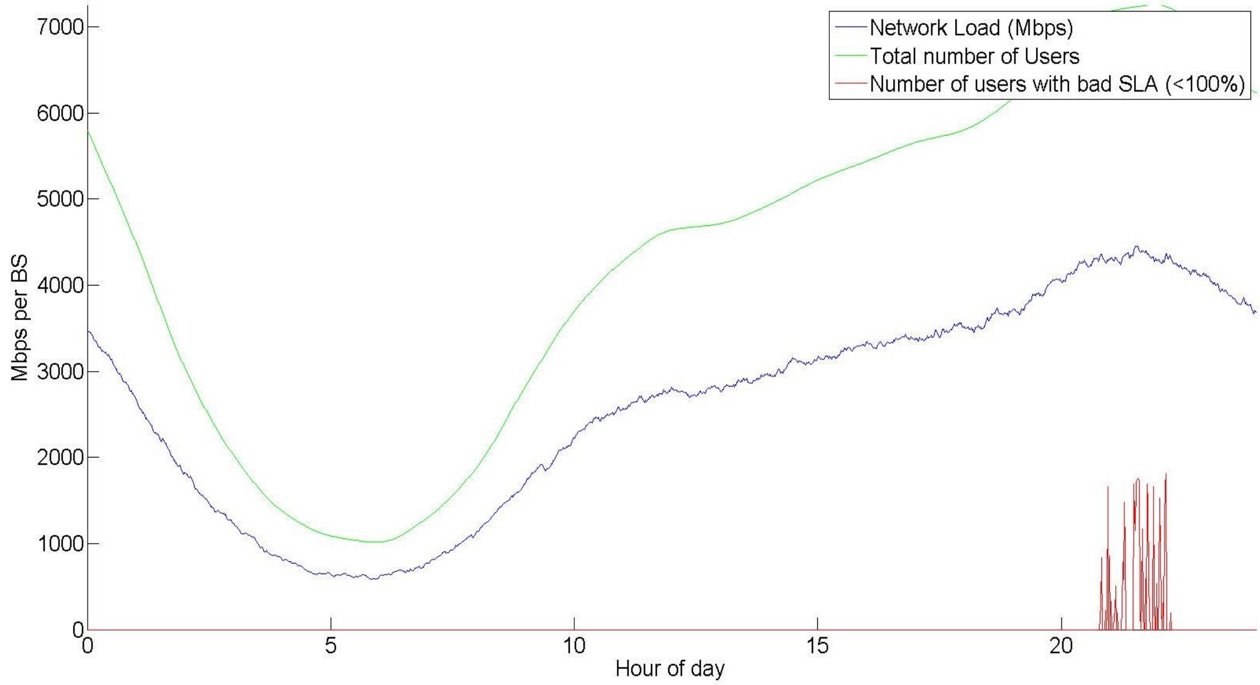


Figure 9 QoS (red curve) of the network using complex policy in realistic scenario.

The results provide consistent and very interesting enhancement results for QoS considerations. Thus, it proves the advantages of the middleware deployment and the sharing of context user data.

3 SEAMLESS OFFLOADING IN HETEROGENEOUS WIRELESS NETWORKS

The number of personal mobile devices as smart phones, tablets, net-books is increasing day by day. These personal mobile devices are all connected to Internet and subscribers expect high data rates when using their mobile devices. Such increase in the number of devices and the subscribers' expectation increases competition among the mobile operators. Therefore, among the all possible solutions, offloading some data traffic to the existing WiFi networks shines brighter. The main purpose of this work is to present some viable solutions to switch between 3GPP data networks and WiFi networks. The QoE (Quality-of-Experience)-based innovative data offloading strategy is also introduced and evaluated in terms of mean opinion score, peak signal-to-noise ratio and communication delay.

3.1 Solution Description

Mobile data traffic has soared drastically in the past few years. The paramount reasons for this are the increasing smartphone usage, several voice and data campaigns, and the trend of watching video streams from different sources including IPTV and online video databases. It is anticipated that the increased interest for data connectivity is likely to put a burden on network capacity in near future. Operators are thus looking for cost-effective solutions to overcome the capacity bottlenecks in 3G infrastructures likely to emerge in high contention traffic scenarios. Several schemes have been offered so far, mostly consisting of temporary solutions to "save the day". They are likely to induce new costs resulting from femtocell or WiMAX, LTE, LTE-Advanced systems deployment.

However, operators realize that such options provide short-term relief only and they require that the target users stay in the operators own network. A more comprehensive solution that addresses real-world user behavior, i.e. that supports network roaming (both in vertical and horizontal directions) and allows users to receive and enjoy high quality services from their operator regardless of their location and choice of network access, is needed.

Such a solution becomes viable through the use of 802.11 technologies, as operators are already expanding their networks with 802.11 technologies such that they can exploit the free-band communication. It is reasonable to expect that the data traffic should be able to offload to operator based WiFi networks, implying vertical handovers between WLAN and 3GPP technologies.

The objective of this work is to investigate the handover solutions in heterogeneous networks and point out the metrics and factors influencing data offloading and related open research issues to the research community. To this extent, in Network Simulator-2 (NS-2) Media Independent Handover (MIH) functionality is utilized as presented by IEEE 802.21 WG [IEEE21] to monitor access networks and to perform a seamless handover execution. IEEE 802.21 is developing standards to enable handover and interoperability between heterogeneous network types including both 802 and non-802 networks. In addition to the utilization of MIH functionality, a multiple attribute decision making (MADM) algorithm based on QoE metrics for decision making phase has been implemented.

In summary, the following major contributions have been made:

(i) User preference as a QoE metric during handover decision making has been embedded into the implemented MADM algorithm.

(ii) Handover execution has been handled both based on QoS and QoE values.

(iii) Based on user experience/subjective measures, the user seamlessly offloaded from 3GPP network to WLAN.

As for decision making functionality, UE or Mobile Network Operator (MNO) selects the access network by considering probabilistic demands. Network related, terminal related, user related and application related metrics need to be considered pertaining to vertical handover decisions. However, the paramount elements amongst them are the user-related ones as QoE is at the very heart of contemporary mobile business performance expectations. Related parameters include throughput, energy consumption of the terminal, security etc. It is interesting to note that an adult's preferences along these dimensions would potentially differ from that of a young person. For instance, security-wise an adult might not prefer to watch videos through Wired Equivalent Privacy (WEP) or WiFi Protected Access (WPA) on WiFi networks but Extensible Authentication Protocol-Subscriber Identity Module (EAP-SIM) on 3GPP network. Maybe this choice could be trivial for a young person and actually he would prefer a free communication band, but considering recently emerging security challenges, operators need to pay importance to the subjects of security and privacy pertinent to each and every user they serve [Hadiji12]. Handoff decision criteria can be categorized as below:

- Network-related: coverage, bandwidth, latency, link quality (RSS (Received Signal Strength), BER (Bit Error Rate), cost, security level.
- Terminal-related: velocity, battery power.
- User-related: user profile and preferences.
- Service-related: service capabilities, QoS, QoE, security level [Hadiji12].

The Quality of Service(QoS) and Quality of Experience(QoE), mobility and network architecture are important factors during decision making or network selection phase. The following QoS and QoE metrics are checked while offloading the data traffic due to the nature of real-time applications:

a) End to end delay (s): This includes processing, queuing in both ingress and egress, and propagation delay. The end-to-end delay of a video signal is the time taken for the packets to enter the transmitter at one end, be encoded into a digital signal, travel through the network, and be regenerated by the receiver at the other end.
b) Data received (Kbps): This is calculated based on the successfully received packets.
c) Packet Loss (%): This is calculated based on the dropped packets due to either network problems or some queuing problems.
d) Throughput (Kbps): this is the total traffic where packets are successfully received by the destination excluding packets for other destinations.
e) MOS Value (Mean Opinion Score): This corresponds to a numerical value, ranging between 1(worst) and 5(best) expressing the quality of the voice telephony or audio perceived by user. It is also used as a QoE metric.
f) Jitter (s): In IP networks, jitter is the variation in the time-of-arrival of consecutive packets. Jitter results from a momentary condition where more packets are trying to get on a particular link than the link can carry away [Logot12].

In network-centric approaches, the goal is often to acquire maximum total allocation in 3GPP and non-3GPP networks while minimizing cost of underutilization and demand rejection.

The handoff decisions are an integral part of mobility management, as terminals (mobile and/or nomadic) crossing through coverage boundaries need to access available networks.

Mobiliy management could be realized either through client-based or network-based mobility protocols.

In 3G networks, GPRS Tunelling Protocol (GTP) is the primary network-based protocol used. It is the protocol which allows end users of a GSM or UMTS network to move from place to place whilst continuing to connect to the Internet as if this user is connected from one location at the Gateway GPRS Support Node (GGSN). It does this by carrying the subscriber's data from the subscriber's current Serving GPRS Support Node (SGSN) to the GGSN which is handling the subscriber's session.

For a client-based mobility protocol, 3GPP Release 8 introduces a client-server based protocol, Dual Stack Mobile IP (DSMIP), to enable seamless handover between 3G and WiFi. DSMIP is a mobility protocol specified in IETF that provides IP address preservation, helping the user to handover freely in IPv4 and IPv6 accesses.

Also, the network architecture of HWNs is an essential factor considering a seamless offloading. Loose coupling is shown in Figure 10 and it features less integration between the two types of networks. In this scenario, the WLAN and cellular networks are two separate access networks.

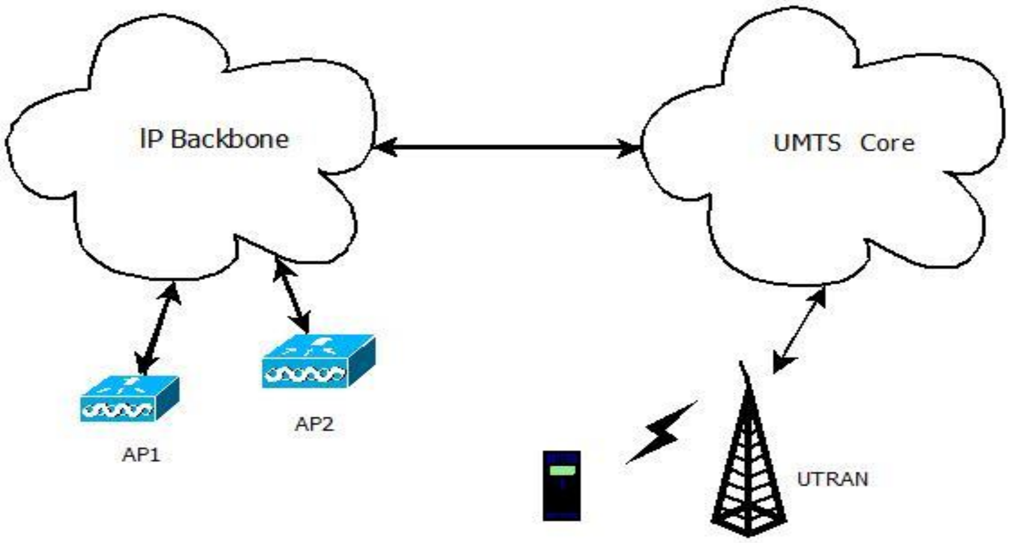


Figure 10 Loose Coupling

Tight coupling illustrated in Figure 12 suggests that WLAN technology is employed as a new radio access technology within the cellular system. Regardless of the access technology, there would only be one common cellular core network [Gang05].

IEEE 802.21 WG was formed to overcome the diversity in the handover mechanisms and to eliminate user-centric methods' drawbacks, and therefore, a common MIHF (Media Independent Handover Function) was introduced. MIHF is an abstraction layer between layer 2 and layer 3. All interfaces on L2 (cellular, WiMax, LTE, WiFi) could communicate with MIHF, and MIHF could transmit the necessary messages to L3 and above such as Session Initiation Protocol (SIP), (Mobile-IP version 4) MIPv4, MIPv6 or Host Identiy Protocol (HIP) [Dutta06].

In IEEE 802.21 framework, there are several factors to determine the handover decision including service continuity, application class, QoS, network discovery and selection, security, power management, and handover policy. In Figure 11, MIHF architecture is illustrated. The most important advantage MIHF brings is not only to provide both L2 and L3 handover but also to allow running make-before-break soft handover mechanism which in return supply us MOS calculation time before the final break execution with the connected access network.

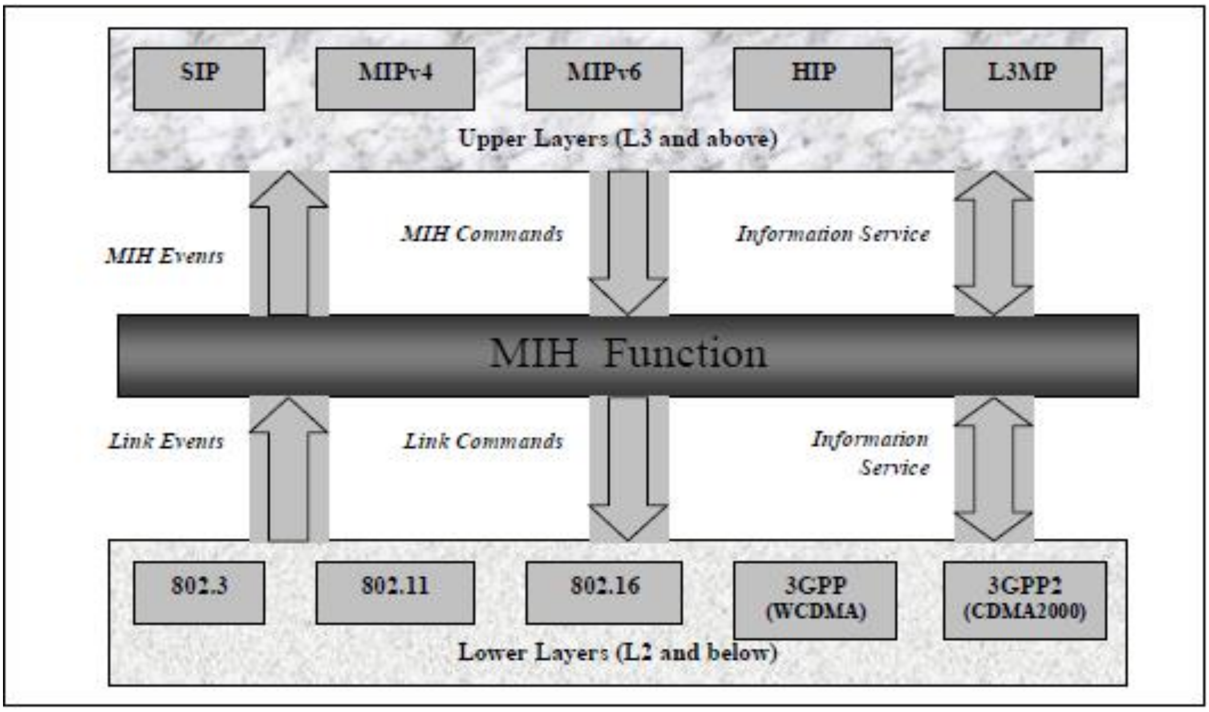


Figure 11 MIHF Architecture

Service delivery and user experience are strongly related items for operators in terms of radio resource management. Technical aspects targeting this issue relate to QoS parameters that can be handled by the platform, at least partially. Subjective psychological issues and human cognitive aspects are typically unconsidered aspects and they directly determine the QoE.

It does not matter how smoothly packets move through your network, if the users find out that services and applications don't meet expectations. Planning must address the factors that underlie QoE for each service that runs on the network, as well as any interactions or inconsistencies between them. For this work, the problems due to the application's packetization and/or encoding/decoding schemes which could also affect QoE are not considered, but only focused on the optimization of the network traffic for the heterogeneous networks.

In this study, TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) [Mark10] is used, due to its easy implementation, as a way of selecting the best target network for a given user's video application. The decision to use this algorithm was made based on the other multiple attribute decision making (MADM) algoritms' performance comparison results. In [Stevens17], four different MADM algorithms (MEW, SAW, GRA, TOPSIS) were evaluated and it was concluded that they all performed very similar. By using this algorithm, the MIH events are triggered such as connect-link or disconnect-link to execute the handover seamlessly. A summary of network selection algorithms [Stevens17] are presented in Table 1.

Table 1 Summary of network selection algorithms

Algorithms and Theories	Purpose	Scheme
TOPSIS	To find the best network based on the selected factors	Collaborative
Stochastic Linear Programming	To find utilization percentage of each access network	Network-centric
AHP and GRA	To rank the access networks in question	User-centric
Game Theory	Bandwidth allocation	Network-centric
Degradation Utility	To find utilization percentage of each access network	Network-centric
Consumer Surplus	To find the best network based on the selected factors	User-centric
Profit Function	To find the best network based on the selected factors	User-centric
Fuzzy Logic Controller	To rank the access networks in question	Collaborative
Objective Function	Bandwidth allocation	Collaborative

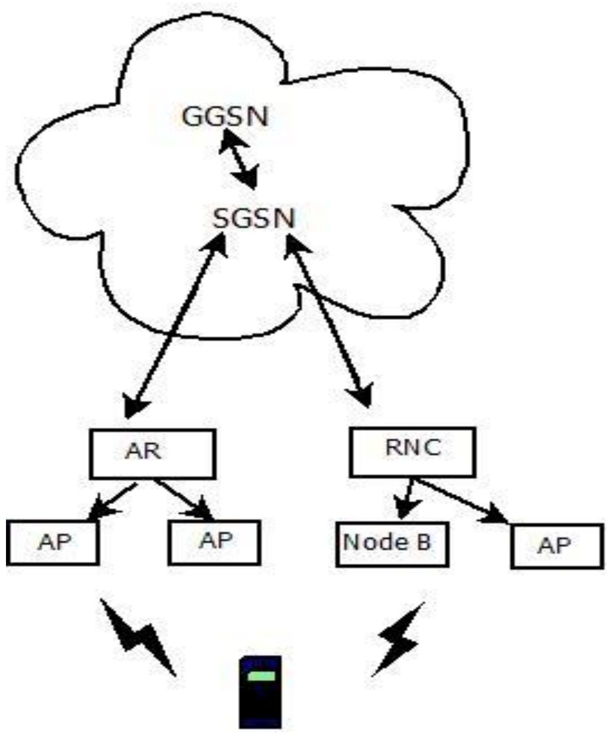


Figure 12 Tight Coupling

3.2 Scenario

Resource management in Heterogeneous Wireless Networks or WiFi integrated cellular networks could be coordinated through user-centric models, network-centric models or collaborative schemes. User centric models offer ease of implementation and scalability, as opposed to the other two approaches, at the expense of reduced overall system efficiency. The network centric models, on the other hand, provide more efficient solutions that improve overall system efficiency (addressing both 3GPP and non-3GPP subsytems), at the cost of increased control overhead and risk of single point of failure. Collaborative solutions, on the other hand, introduce complexity; however, in return, offer drastic performance difference with respect to network-centric solutions in terms of QoE. Basically, in a collaborative solution, UE data such as Received Signal Strength (RSS) or Channel Quality Indicator (CQI) along with access network metrics obtained from an operator server are combined in decision making phase. Then, based on the implementation choice the decision is executed either on network-side or user-side. Figure 13 illustrates different access technologies' usage for a collaborative solution where UE executes the decision by integrating a MIIS (Media Independent Information Service Server) to utilize the MIH functionality. Table 2 presents a summary of some media access technology characteristics.

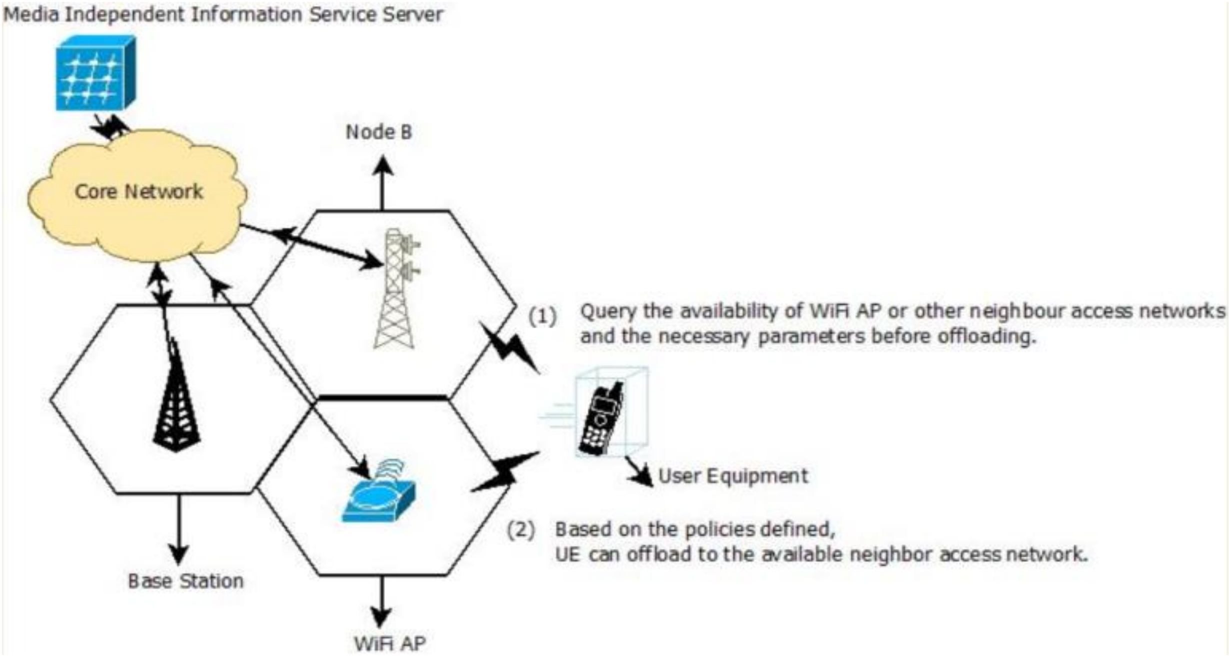


Figure 13 A MIIS adapted handover solution

Table 2 Comparison of access networks

Class	Technology	Data Rate	Range
Cellular	2G,2.5G,3G,3.5G,3.9G,4G	20Kbps< x <110Mbps	Cellular coverage
WLAN	802.11a 802.11b 802.11g 802.11n	54 Mbps 11 Mbps 54 Mbps <600 Mbps	100 m
WPAN	802.15 Bluetooth	0.1-1 Mbps	<10 m
WMAN	802.16 WiMAX	70 Mbps	50 km

Base functions that are common to the above mentioned inter-technology offloading or handover mechanisms are: (i) Resource Monitoring including network discovery, (ii) decision making including network selection and (iii) decision enforcement [Piamrat11], [Gang05].

Figure 14 demonstrates a generic approach for resource management in heterogeneous wireless networks [Piamrat11].

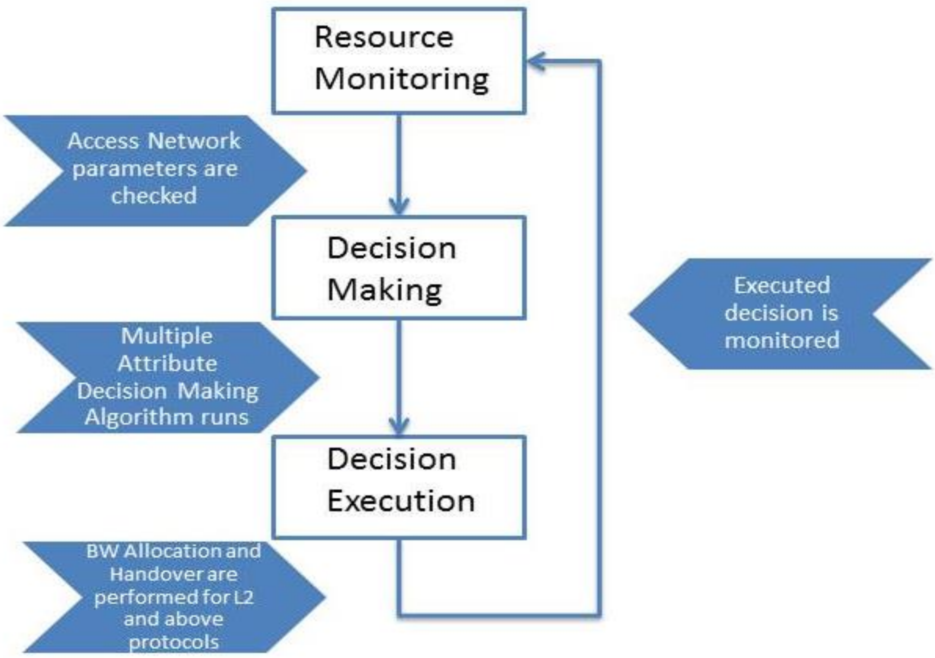


Figure 14 A generic approach for resource management

Resource monitoring helps decision maker identify and discover access networks available along with the corresponding QoS (Quality of Service) and/or QoE indicators. Depending on the approach utilized, the decision maker could be the user-equipment (UE) or network, meaning either each cell will broadcast its connection information to UEs or UE will retrieve the cell information from both 3GPP and non-3GPP networks. In 4G networks, access network discovery could be controlled by Evolved Packet Core (EPC) along with access network discovery and selection function (ANDSF). However, 3G networks are missing such a core system and require either an additional device or an additive functionality to the existing serving nodes in the infrastructure or a user-centric approach. 3GPP or trusted/untrusted non-3GPP (WiFi) networks could be discovered and monitored by this functionality [Bennis13] [Fadel12].

The handoff decision algorithm aims at selecting a network for a particular service that can satisfy objectives based on some criteria (such as low cost, good RSS, high MOS, optimum bandwidth, low network latency, high reliability and long life battery) and taking into account the preferred access network of user. Some techniques used for network-centric solutions such as stochastic programming, game theory and utility function could be performed in this respect and are explained as below [Piamrat11].

Stochastic linear programming obtains maximum allocation in each network by using probabilities related to allocation, underutilization, and rejection in Heterogeneous Wireless Networks (HWNs). Game theoretic approaches take advantage of the bankruptcy game, and efficient bandwidth allocation and admission control algorithms are developed by utilizing available bandwidth in each network. In utility function, operator prioritizes users and classifies services to allocate bandwidth for the users [Taha04] [Niyato06] [Yang05].

In user-centric solutions, the users themselves (or their agents) make the decisions, often prioritizing the needs and objectives of the individual users. Analytical hierarchy processes help ranking the networks based on induced QoS indicators, by checking user’s requirements and network conditions. Proposed approaches make use of the consumer surplus model and similar economic theory based techniques. Users are often modeled to have profit functions amounting to the difference between bandwidth gain and handoff cost for each network is computed. The most appropriate network is found through utility maximization [Song05] [Ormond05] [Liu06].

As for the collaborative models, fuzzy logic controller ranks the candidate networks based on the user’s selection criteria, network data rate and Signal-to-Noise Ratio (SNR). In objective function, user’s RSS, network’s queue delay and policy preferences such as cost are fed as input parameters, and the function provides the allocation of services to APs and terminals. Lastly, in TOPSIS, the best path for flow distribution on muti-homed end-hosts is computed. Also, network’s QoS (delay, jitter, and BER), user’s traffic class and most importantly QoE are also considered [Wilson05] [Koun07] [Nacef08].

Other options could be to harness impatient or patient algorithms which are based on user-centric solutions. The Impatient algorithm uses a very simple policy: use 3G whenever WiFi is unavailable; else use WiFi. The Patient waits and sends data on WiFi until the delay tolerance threshold, and only switches to 3G if all of the data are not sent on WiFi before the delay tolerance threshold [Bala10].

In the decision enforcement phase, the decisions obtained are executed as its name implies. Basically, the necessary bandwidth is allocated for the user. In user-centric approach, the obtained decisions are not always achieved if the network does not accept the request; contrary to network-centric approach where solutions are always achieved since it is the network operator who controls all the resources.

The user-centric methods do not consider the possibility that users could harness the network based on their own profit so that congestion could increase whereas in a network-centric solution operator could decide and offload the traffic based on the general state of the infrastructure. Furthermore, power consumption for user equipment is not considered in user-centric methods.

Some other problems related to user-centric and network-centric models are: a device would connect at Layer 2, but not at network layer. Also, for user-centric schemes UEs would connect one of APs available based only on signal strength, and end-up with wrong assignment such as application class or QoS requirement are not met. Adding the increasing number of interfaces such as WiMAX, WiFi and cellular network, the burden of UE would extend to cover multiple interfaces.

IEEE 802.21 WG was formed to overcome the problems related to inter-layer communication in the handover mechanisms and to eliminate user-centric and network-centric methods’ drawbacks, and therefore, a common MIHF (Media Independent Handover Function) was introduced. MIHF is an abstraction layer between layer 2 and layer 3. All interfaces on L2 (cellular, WiMax, LTE, WiFi) could communicate with MIHF, and MIHF could transmit the necessary messages to layer 3 and above such as SIP, MIPv4, MIPv6 or HIP [Dutta06].

3.3 Results

For the simulation, NS 2.29 simulator integrated with EURANE (Enhanced UMTS Radio Access Network Extension), NIST (National Institute of Standards and Technology) and EVALVID packages is used to evaluate the video performances in a heterogeneous network during a handover execution where the TOPSIS algorithm results were utilized. For video streaming, channel utilization and MOS of service should be of utmost importance. Decision parameters of TOPSIS are as following:

i. MOS: Mean Opinion Score is considered as a subjective measure. Currently, it is more often used to refer to one or another objective approximation of subjective MOS. Although all "MOS" metrics are intended to quantify QoE performance and they all look very similar (values between one and five with one or two decimal places), the various metrics are not directly comparable to one another. ITU P.800 and P.830 define the MOS scale as showed in Table 3.

ii. PSNR (dB): The peak signal-to-noise ratio is used as an objective measurement of the restored image quality. PSNR is most commonly used to measure the quality of reconstruction of lossy compression codecs which is in this case MPEG-4.

$$PSNR = 20 \log \frac{Vpeak}{MSE}$$

Vpeak = $2^k - 1$ where k is equal to number of bits per pixel (luminance component). MSE is the mean squared error.

iii. CQI: Channel quality indicator is reported by UE and is calculated using BLER and SNR values. It is a vital parameter to estimate the UMTS air interface quality. The UE type that is assumed in the simulator is 3GPP UE category 1 to 6. In the simulation, the highest CQI value was accepted as 22. However, it varies between 1 and 22.

iv. QoS: Quality of service level of the access point (AP) is utilized in the algorithm to determine the link-quality of WiFi network. Voice = Platinum = 6, Video = Gold = 5, Best Effort = Silver = 3, Background = Bronze = 1

v. Security Policy used in WiFi network: WPA or WPA2 cannot be used for a seamless and operator controlled solution. EAP-SIM is required to do so.

vi. Channel Utilization: It is a WiFi network parameter, and is monitored for a stable traffic level and to prevent under or over utilization. The range is defined in percentages (0%-100%).

vii. Client SNR: Measured signal-to-noise ratio is a critical and widely used metric to obtain the experienced WiFi quality per user.

viii. User Preference: For a businessman security and quality level could be extremely important whereas for a student the cost is of the utmost importance.

Table 3 ITU-R Quality and Impairment Scale

Scale Impairment	**Quality**	**Impairment**
5	Excellent	Imperceptible
4	Good	Perceptible, but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

QoE assessment could be performed with subjective tests with humans, but by using this scheme a handover execution in real-time cannot be made. Other approaches to the problem of QoE assessment includes utilizing objective testing to predict the MOS value of a service.

These solutions need original signals (for real time applications e.g., ITU-T objective measurement standards like PESQ (P.862), E-model (G.107) etc.) and are computationally complex [Rec01], [Rec03]. Therefore, the PSNR frame by frame is calculated and maped into the corresponding MOS value as in Table 4.

Table 4 PSNR to MOS mapping

PSNR [dB]	**MOS**
> 37	5 (Excellent)
31 - 37	4 (Good)
25 - 31	3 (Fair)
20 - 25	2 (Poor)
< 20	1 (Bad)

When simulating the heterogeneous network in tight-coupling architecture, a case where a video is downloaded in the beginning is assumed. The user was connected to an UMTS network and in this network throughput was 45Kb/s which is not even acceptable for voice networks since 64Kb is used for bearer payload whereas 16Kb needs to be used for signaling purposes in each direction, making a total of 80Kb/s at least for a good quality voice traffic. Considering the traffic, video requires a lot more throughput for an acceptable communication.

One UE's speed was 1m/s, and after 2 seconds where video is transmitted with a rate of 30fps, the MIH module discovered the WLAN network and requested the target WLAN network's metrics regarding QoS (jitter, delay, packet loss) and also user preference such as security (i.e. EAP-SIM or WEP) along with MOS. Based on the TOPSIS algorithm, user was attached to the target WLAN network by using MIH functionality at frame of 60. Between frames 60 and 300, user experienced MOS values between four and five.

In Figure 15, cumulative distribution function (CDF) of the delay has been presented for high, moderate, and slow UMTS throughput rates which are 672 Kb/s, 340Kb/s, and 45Kb/s respectively. Lost frames acquire a delay of 0. Thus, the start of the CDF-lines is the percentage of lost frames.

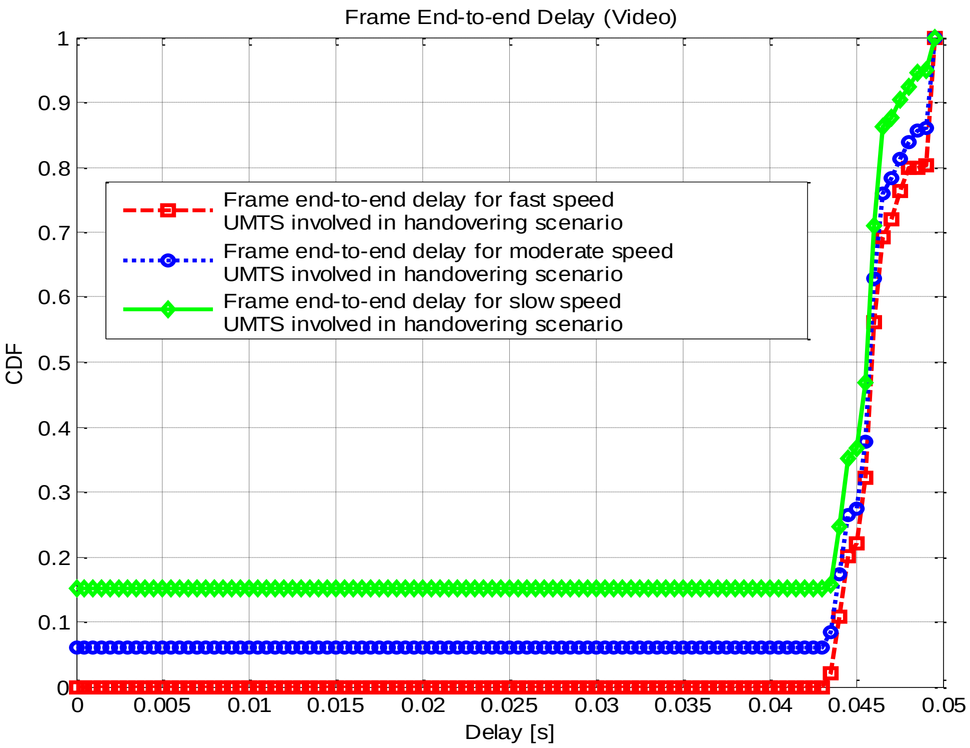


Figure 15 Frame End-to-end Delay

PSNR vs Frames is presented in Figure 16. It is observed that low throughput and high mobility affect PSNR severely and both of these attributes are in close correlation. After frame 60, the network handovers the user to an available WiFi AP based on the implemented MADM algorithm, and consequently, PSNR value and user experience increases.

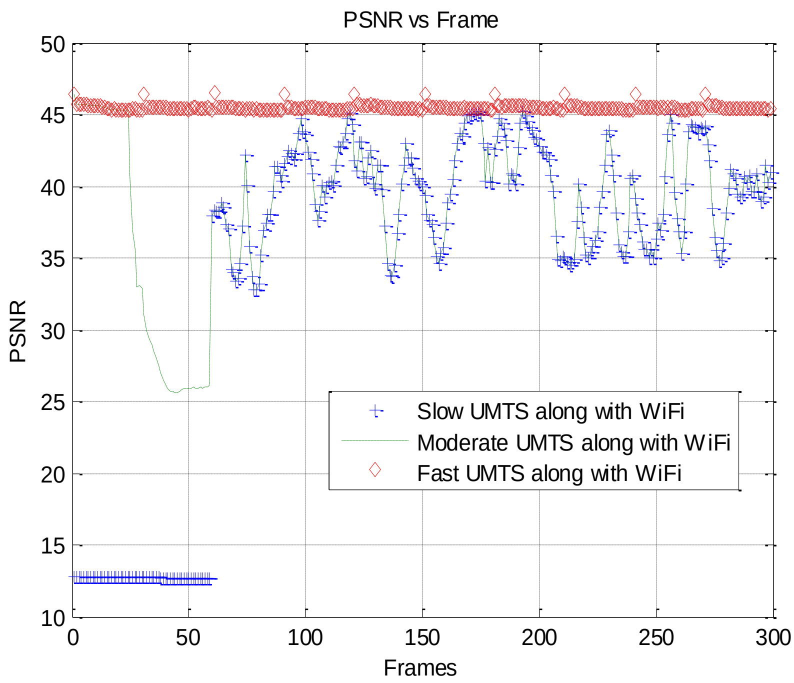
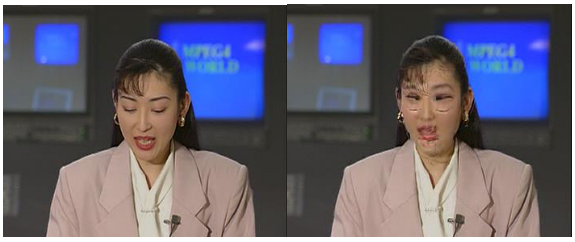
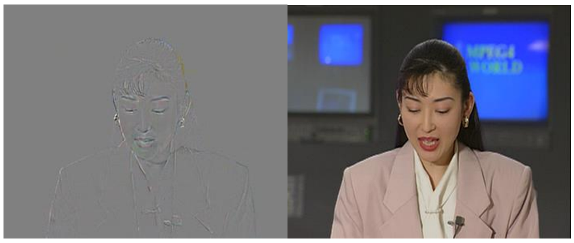


Figure 16 PSNR vs Frames

In Figure 17, the received frames of 59 for the three different PSNR obtained just before the handover executed to WiFi are compared. In three scenarios, WiFi traffic started after frame 60, and WiFi metrics in terms of QoS were always superior to UMTS; also, user preferred high MOS value with low cost. Basically, it is shown that for high throughput scenarios as in Figure 17(a), received frame quality is high and network does not have to offload the user in terms of user experience. However, for moderate and low throughput scenarios presented in Figure 17(b) and Figure 17(c) respectively, network operator has to make a decision to offload the user to an available WiFi AP so that the both the burden on the 3GPP AP could be lessened and the user could experience better quality videos.



(a) **(b)**



(c) **(d)**

Figure 17 (a) Received CIF resolution frame in MPEG-4 XviD for fast UMTS – PSNR:45.39dB (b) Received CIF resolution frame in MPEG-4 XviD for moderate UMTS – PSNR:26.14dB (c) Received CIF resolution frame in MPEG-4 XviD for slow UMTS – PSNR:12.74dB (d) Transmitted original CIF resoluton frame in MPEG-4 XviD

4 CAPACITY AWARE MULTI-USER LOAD BALANCING FOR HETEROGENEOUS NETWORKS FOR MULTIUSER SCENARIOS

The main purpose of this work is to balance the load between 3GPP (LTE) and WLAN networks deployed within an integrated architecture. A novel capacity-aware multi-user multi-attribute decision making algorithm is presented and evaluated in terms of mobile user distribution and total channel utilization in the heterogeneous network. The proposed algorithm is shown to enhance total channel utilization of heterogeneous networks compared to standard single-user decision making algorithms.

4.1 Solution Description

The objective of this work is to investigate handover decision making algorithms in heterogeneous wireless networks and point out the metrics and factors influencing data offloading and related open research issues to the research community. To this extent, a capacity aware multi-user multiple attribute decision making (MADM) algorithms based on Quality of Experience (QoE) metrics has been developed and evaluated. The proposed capacity aware multi-user load balancing algorithm optimizes total benefit of the system that is balanced according to total channel utilization among different heterogeneous wireless networks. The proposed algorithm is shown to enhance total channel utilization of heterogeneous networks compared to standard single-user decision making algorithms.

4.1.1 Multi-user Offloading Algorithms For Hetereogeneous Networks

In this study, a multiple user multiple attribute decision making (MADM) problem targeted for heterogeneous network access within an integrated mobile network architecture is considered. For the system model, the following conditions are assumed to represent the MADM problem.

- The total users set in the system is denoted as $U = \{u_1, u_2, u_3, ..., u_k\}$ where k (k>=2) denotes number of users.
- The multiple users' set involved in the decision making process are denoted as $V = \{v_1, v_2, v_3, ..., v_{k'}\}$ where k' (k'<=k) denotes number of users under multiple coverage.
- The multiple attribute set is denoted as $S = \{s_1, s_2, s_3, ..., s_m\}$ where m (m>=2) denotes number of possible attributes.
- The multiple decision point set is denoted as $E = \{e_1, e_2, e_3, \ldots, e_P\}$ where there are p (p≥2) possible decision points.

The weight set is denoted as $w = \{w_1, w_2, w_3, \ldots, w_m\}$, where each weight w_i is the weight assigned to attribute s_i $i \in \{1,2,\ldots,m\}$. In this study, TOPSIS is used as the core algorithm [Markovic10], due to its easy implementation, as a way of selecting the best target network for a set of given users. The decision to use this algorithm was made based on the other multiple attribute decision making (MADM) algorithms' performance comparison results. In [Stevens17], four different MADM algorithms (MEW, SAW, GRA, TOPSIS) were evaluated and it was concluded that they all performed very similar.

4.1.1.1 TOPSIS

TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) [Mark10], due to its easy implementation, is a suitable candidate to select the optimal target network for a given a set of given observed attributes for a user.

In the first step of TOPSIS algorithm a decision matrix **A** is created:

$$\mathbf{A} = \left[a_{ij} \right] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{p1} & a_{p2} & \dots & a_{pm} \end{bmatrix} (i = 1, \dots., p; j = 1, \dots., m)$$

In matrix **A**, m refers to size of the multiple attribute set such as link quality, MOS of the target network for the given application, user preference (cost security), etc and *p* refers to size of the multiple decision point set decision points target networks which can be LTE, WLAN or D2D (device-to-device). Note that that all the attributes are transformed to have positive impact if necessary.

In second step, a normalized decision matrix is formed by using the following equation:

$$r_{ij} = \frac{a_{ij}}{\sqrt{\sum_{k=1}^{p} a_{kj}^2}}$$

Then the normalize matrix R is obtained as:

$$\mathbf{R} = [r_{ij}] = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ r_{21} & r_{22} & \dots & r_{2m} \\ . & & & . \\ . & & & . \\ . & & & . \\ r_{p1} & r_{p2} & \dots & r_{pm} \end{bmatrix}$$

In third step, a weighted normalized decision matrix is created by multiplying each column of the matrix by corresponding weight w_i where $\sum_{i=1}^{m} w_i = 1$ by using the following equation:

$$\mathbf{v_i} = w_i * \mathbf{r_i}, \qquad \mathbf{r_i} = [r_{1i}, \dots., r_{pi}]^T, \qquad i = \{1, 2, \dots, m\}$$

In fourth step, the positive (A^*) and negative (A^-) solutions are formed by using the following formulas:

$$A^* = \left\{ (\max_i v_{ij} \,|\, j \in \{1, 2, ..m\}) \right\}$$

$$A^- = \left\{ (\min_i v_{ij} \,|\, j \in \{1, 2, ...m\}) \right\}$$

At the end of fourth step, sets are formed as $A^* = \{v_1^*, v_2^*, \dots, v_n^*\}$ and $A^- = \{v_1^-, v_2^-, \dots, v_n^-\}$.

By calculating the Euclidean distance S_i^* of each multiple decision point from the positive point A^* and S_i^- of each multiple decision point from the negative point A^-.

$$S_i^* = \sqrt{\sum_{j=1}^{p} (v_{ij} - v_j^*)^2}, \quad i = \{1, \dots, p\}$$

$$S_i^- = \sqrt{\sum_{j=1}^{p} (v_{ij} - v_j^-)^2}, \quad i = \{1,..., p\}$$

In the final step, the relative similarity of the alternatives from the positive and negative points is calculated as:

$$C_i = \frac{S_i^-}{S_i^- + S_i^*}, \quad i = \{1,..., p\}$$

where $0 \le C_i \le 1$ the final solution is selected by:

$$e^* = e_{i^*} \text{ where } i^* = \arg\max_i C_i, \quad i = \{1,..., p\}$$

Multiple attribute sets are used as provided in Section 3.3. In the next two sub-sections, two algorithms are defined. First algorithm developed is a Multi-user TOPSIS with capacity-aware characteristic where channel utilization parameter is of the utmost importance for the 3GPP network to balance the channel allocations. With this type of multi user algorithm, the total system benefit is considered as important. Second algorithm is Standard TOPSIS (ST) algorithm. With this method each user's individual benefits are considered individually as they arrive.

4.1.1.2 Capacity aware multi-user iterative TOPSIS (CAT) algorithm

In order to obtain certain benefits for access channel selection and resource allocation problem between multiple users, Capacity aware iterative multi-user TOPSIS algorithm is proposed.

Input: Set of technology E, total channel utilization threshold for each technology CU_{th}^e, and the TOPSIS matrix of user $v_{i'} \in V$ denoted by

$$\mathbf{A}^{i'} = \left[a_{ij} \right] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{p1} & a_{p2} & \dots & a_{pm} \end{bmatrix} (i = 1,...., p; j = 1,...., m)$$

Output: Capacity-aware channel utilization vector $\mathbf{CU^e} = [CU_1^e, CU_2^e, \dots, CU_{k'}^e]$, e ∈ E.

Step1: Set $\mathbf{CU^e}$=[0] and i'=0 (i' ≤ k' is the user number)

Step2: Put i'= i'+1, as user $v_{i'}$ arrives.

Step3: Run TOPSIS algorithm using $A^{i'}$ and select the optimal decision point $e^* = e_n \in E$ and construct coincidence coefficient

$$\delta_{i'}^{e} = \begin{cases} 1 \; if \; e = e* \\ 0, otherwise \end{cases} \quad , \forall \, e \in \mathrm{E}$$

Step4: Update the temporary channel utilization vector $\widehat{\mathbf{CU}}^{e*} = \mathbf{CU^{e*}}$ and put $\widehat{\mathrm{CU}}_{i'}^{e*} = a_{nc}$ where c denotes the column number in $A^{i'}$ for attribute corresponding to CU ∈ S.

Step5:

- If $(\sum_{j=1}^{i'} \delta_j^{e*} CU_j^{e*} \le CU_{th}^{e*})$
    - $\mathbf{CU^{e*}} = \widehat{\mathbf{CU}}^{e*}$
- else
    - E = E \ e*
        - If E = {} then e* = WLAN

- else go to **Step 3**

Using this multi user algorithm, the total system benefit is considered as the first criteria to optimize and the minimum bit-rates are maximized by assigning the users in the intersection area to WLAN where 3GPP utilization is high. In the analysis, channel utilization is selected as the most important parameter for the 3GPP network to balance the channel allocations.

4.1.1.3 Standard TOPSIS (ST) method

With standard TOPSIS method, user's individual's benefits are considered. The TOPSIS algorithm is explained in [Mark10]. The method details are explained in the following steps:

Input: Set of technology E, and the TOPSIS matrix of user $v_{i'} \in V$ denoted by

$$\mathbf{A}^{i'} = \left[a_{ij}\right] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ . & . & \dots & . \\ . & . & \dots & . \\ . & . & \dots & . \\ a_{p1} & a_{p2} & \dots & a_{pm} \end{bmatrix} (i = 1, ..., p; j = 1, ..., m)$$

Output: Standard TOPSIS channel utilization vector $\mathbf{CU^e} = [CU_1^e, CU_2^e, ... , CU_{k'}^e]$, $e \in E$.

Step1: Set $\mathbf{CU^e}$=[0] and i'=0 (i' ≤ k' is the user number)

Step2: Put i'= i'+1, as user $v_{i'}$ arrives.

Step3: Run TOPSIS algorithm using $A^{i'}$ and select the optimal decision point $e^* = e_n \in E$ and construct coincidence coefficient

$$\delta_{i'}^{e} = \begin{cases} 1 \; if \; e = e* \\ 0, otherwise \end{cases} \quad , \forall \, e \in \mathrm{E}$$

Step4: Update the channel utilization vector by $CU_{i'}^{e*} = a_{nc}$ where c denotes the column number in $A^{i'}$ for attribute corresponding to CU ∈ S.

4.2 Scenario

As can be seen in Figure 18 the set of users that are in the coverage area of both WLAN and LTE are within the shaded area. These users have the high potential of handover and have to make a smart decision to select the best access point. Therefore, the method runs on the scenarios based on the users that are concentrated on this region.

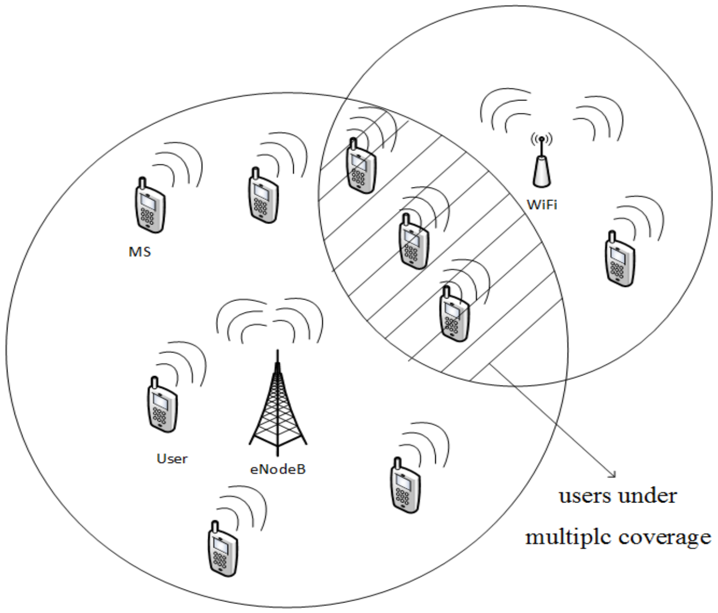


Figure 18 A sample user distribution map under multiple wireless technology coverage

4.3 Results

In this section, simulation scenario and the simulation results for the algorithms provided in Section 4.1 are presented.

4.3.1 Simulation Scenario

For simulations, a video streaming use case scenario is selected. In order to offload a video streaming seamlessly, only most relevant parameters were selected such as Channel utilization, MOS, QoS, delay, and energy as shown in Table 5.

Table 5 Network Selection Criterias for Video Streaming

Attribute Set	S	**Weight (w_i)**
MOS	s_2	0.25
QoS	s_3	0.1
Energy	s_5	0.05
Channel Utilization	s_1	0.5
Delay	s_4	0.1

For this scenario, channel utilization and MOS of service are of utmost importance and therefore the weight coefficients are distributed accordingly as shown in Table 5.

For the purpose of comparison, the same attribute values and the same weights are assigned to attributes for the different above algorithms. Note also that assignment of weights could be initiated by either user or operator or collaboratively.

In order to compare algorithms, an environment where k'= 4 users are under multiple coverage and their respective attribute weight values are same is created. The decision points are wireless technologies where users handover to will be limited to WLAN and 3GPP networks, i.e. E = {WLAN, LTE} and

$$\mathbf{A}^1 = \begin{bmatrix} 6 & 3 & 5 & 6 & 7 \\ 4 & 3 & 4 & 6 & 3 \end{bmatrix}, \quad \mathbf{A}^2 = \begin{bmatrix} 1 & 5 & 5 & 4 & 6 \\ 4 & 3 & 3 & 7 & 6 \end{bmatrix}, \quad \mathbf{A}^3 = \begin{bmatrix} 1 & 5 & 5 & 7 & 6 \\ 7 & 1 & 2 & 2 & 1 \end{bmatrix},$$

$$\mathbf{A}^4 = \begin{bmatrix} 1 & 5 & 5 & 2 & 6 \\ 7 & 6 & 6 & 4 & 4 \end{bmatrix}$$

The channel utilization threshold for LTE is CU_{th}^{LTE} = 8 units, the channel utilization threshold for WLAN is CU_{th}^{WLAN} = 12 units.

4.3.2 Performance Evaluations

The total user distributions on a HetNET comprising 3GPP (LTE Network) and WLAN access networks for CAT, ST algorithms as well as ALL 3GPP and ALL WLAN scenarios is shown in Table 6 . When CAT algorithm is applied, the number of users among different access technologies is %25 and %75 for 3GPP and WLAN respectively. For the ST algorithm, the distributions become %75 and %25 for 3GPP and WLAN respectively.

Table 6 Total User Distribution and Channel Utilizations (%) of All Algorithms

	USERS DISTRIBUTION		**TOTAL CHANNEL UTILIZATION (%)**	
	3GPP (LTE)	**WLAN**	**3GPP (LTE)**	**WLAN**
CAT	%25	%75	%50	%66
ST	%75	%25	%225	%50
ALL 3GPP	%100	%0	%275	0
ALL WLAN	%0	%100	0	%75

Similarly, the total channel utilization distributions for CAT, ST algorithms as well as ALL 3GPP and ALL WLAN scenarios are also shown in Table 6 . The total channel utilization percentages are calculated by dividing sum of the demands of users for channel utilization over channel utilization thresholds of each technology.

When the CAT algorithm is applied, where total benefit of the system is optimized according to multiple attributes described above, balancing the total channel utilizations among 3GPP (LTE) and WLAN technologies provides lower channel utilizations. The CAT algorithm yields the total channel utilization percentages of %50 and %66 for 3GPP and WLAN Technologies respectively. On the other hand, when ST algorithm is applied, TOPSIS algorithm will prioritize individual user benefits, namely individual QoE. It is clearly seen that ST algorithm could lead to high channel utilization which consequently would decrease MOS substantially for the corresponding access networks. For ST algorithm, channel utilization percentage of %225 represents over channel utilization for 3GPP (LTE).

The important point to notice for ST algorithm is that even though the expected individual QoE will be high with this type of algorithm, due to over-allocation in one access network after the handover decisions are executed, the users will suffer from either ping-pong effect or real-time network changes which will induce additional burden into the system both in terms of network and terminal. However with CAT algorithm, after prioritizing channel utilization and MOS attributes, channel utilizations are optimized between 3GPP and WLAN access networks, which in return increases the QoE of users compared to simple ST algorithm.

From operator point of view, CAT algorithm works best in terms of channel utilization or load balancing; however, with this type of scheme some attributes (other than channel utilization) observed by users can be diminished compared to ST algorithm.

Finally, in terms of total channel utilization, CAT and ST are compared with ALL 3GPP and ALL WLAN scenarios where no algorithm is implemented and all users are either on 3GPP or WLAN networks. One can observe that channel utilizations for ALL 3GPP exceeds channel utilization thresholds which in return will over-allocate the system, adding additional burden to the operators of these wireless access technologies.

5 Joint Offloading And Scheduling Strategies For dual Mode Small Cells

The deployment of small cell base stations, SCBSs, overlaid on existing macrocellular systems is seen as a key solution for offloading traffic, optimizing coverage, and boosting the capacity of future cellular wireless systems. The next generation of SCBSs is envisioned to be multimode (i.e. capable of transmitting simultaneously on both licensed and unlicensed bands). This constitutes a cost-effective integration of both WiFi and cellular radio access technologies that can efficiently cope with peak wireless data traffic and heterogeneous quality of service requirements. To leverage the advantage of such multimode SCBSs, the novel proposed paradigm of cross-system learning is discussed by means of which SCBSs self-organize and autonomously steer their traffic flows across different RATs. Cross-system learning allows the SCBSs to leverage the advantage of both the WiFi and cellular worlds. For example, the SCBSs can offload delay-tolerant data traffic to WiFi, while simultaneously learning the probability distribution function of their transmission strategy over the licensed cellular band. The basic building blocks of cross-system learning are first introduced and then a preliminary performance evaluation in a Long-Term Evolution simulator overlaid with WiFi hotspots is provided. Remarkably, it is shown that the proposed cross-system learning approach significantly outperforms a number of benchmark traffic steering policies.

5.1 Solution Description

In the context of cellular and WiFi integration, the goal of every SCBS is to devise an intelligent and online learning mechanism to optimize its licensed spectrum transmission, and at the same time leverage WiFi by offloading delay-tolerant traffic. The developed procedure, dubbed cross-system learning, is rooted in the fact that every small cell optimizes its long-term performance metric, as a function of its traffic load, interference levels, and users' heterogeneous traffic requirements. In addition, unlike standard reinforcement learning (RL), the cross-system learning procedure allows players to implicitly coordinate their transmissions with no information exchange, as well as to leverage the coupling between LTE and WiFi, which, as shown below, increases the overall network performance and significantly speeds up the convergence. The cross-system learning framework is composed of the following interrelated components. Subband selection, power level allocation, and cell range expansion bias: Every SCBS learns over time how to select appropriate subbands with their corresponding transmit power levels in both licensed and unlicensed spectra, in which delay-tolerant traffic is steered toward the unlicensed spectrum. In addition, every SCBS learns its optimal CRE bias to offload the macrocell traffic to smaller cells. Traffic-aware scheduling: Once the small cell acquires its subband, the scheduling decision is traffic-aware, taking into account users' heterogeneous QoS requirements (throughput, delay tolerance, and latency).

During cross-system learning, every SCBS minimizes over time its regret of selecting strategies yielding lower payoffs, while experimenting with other strategies to improve its long-term utility estimation. The considered behavioral assumption is that small cells are interested in choosing a probability distribution over their transmission strategies that minimizes the regret, where the regret of SCBS k for not having played action $q_k^{(l,s,b)}$ from $n = 1$ up to time t is defined as

$$r_{k,q_k^{(l,s,b)}}(t) = \frac{1}{t}\sum_{n=1}^{t} \hat{u}_k\left(q_k^{(l,s,b)}, p_{-k}(n)\right) - \hat{u}_k(n)$$

where $u_k(n)$ is the instantaneous utility observation (i.e. feedback) of SCBS k at time n, obtained by constantly changing its strategy. In addition, to calculate its regret, every SCBS k estimates its utility function $\hat{u}_k(.)$ when taking a given action based on local information. The rationale is as follows: If the regret is strictly positive, SCBS k would have obtained a higher average utility by playing action $q_k^{(l,s,b)}$ during all previous time instants; thus, SCBS k

"regrets" not having done so. In contrast, if the regret is negative, SCBS k does not regret its strategy selection. Therefore, each SCBS needs to strike a balance between choosing actions that yield lower regrets (more often than those with higher regrets) and playing any of the other actions with a non-zero probability. The behavioral rule of every SCBS can be modeled by the following probability distribution subject to the maximum transmit power constraints where

$$\beta_k(\boldsymbol{r}_k^+(t)) \in \arg\min_{\pi_k}\left[\sum_{\boldsymbol{p}_k\in A_k}\pi_{k,\boldsymbol{p}_k} r_{k,\boldsymbol{p}_k}(t) + \frac{1}{\kappa_k}H(\pi_k)\right]$$

where $r_k^+ = \max(0, r_k(t))$ denotes the vector of positive regrets, and $H(.)$ represents the Shannon entropy function of the mixed strategy p_k. The temperature parameter κ_k represents the interest of SCBS k in choosing other actions. The unique solution to the right side of the continuous and strictly convex optimization problem

$$\beta_{k,q_k^{(l,s,b)}}(\boldsymbol{r}_k^+(t)) = \frac{\exp\left(\kappa_k r^+_{k,q_k^{(l,s,b)}}(t)\right)}{\sum_{\boldsymbol{p}_k\in A_k}\exp\left(\kappa_k r^+_{k,\boldsymbol{p}_k}(t)\right)}$$

Furthermore, given users' different QoS requirements, the cross-system learning framework leverages WiFi, in which the learning process carried out over WiFi is faster (from a timescale perspective) than that on the cellular band. More concretely, inspired by the well-known turbo principle, the output (i.e., feedback) from the WiFi learning process is used to update the cellular learning process. As shown later, this notion of timescale significantly reduces the convergence time of the traffic steering algorithm compared to standard RL, and improves the overall performance.

Once the SCBSs select their subbands using cross-system learning, they engage in a proactive and traffic-aware scheduling procedure on the selected subband's resource blocks. The scheduling algorithm is proactive and traffic-aware in nature as it incorporates users' traffic requirements. Notably, the scheduling decision is not only based on the instantaneous channel condition but also on the completion time (delay) and service class of each transmission. For that, within every small cell, all users are sorted in ascending order as a ratio of their remaining file size and estimated average data rate. Then SCBS k computes a metric $D_{k_i}(t)$, which is a function of the position of UE k_i and the number of UEs served by SCBS k at time t. Finally, UE k_i^* is scheduled such that:

$$k_i^* = \arg\min_{k_i} D_{k_i}(t)$$

5.2 Scenario

The scenario looks into the problem of dual mode small cells transmitting simultaneously on LTE and WiFi bands. The problem boils down to designing self-organizing load balancing solutions that optimally balances the load among both RATs. The proposed approach is based on reinforcement learning techniques adapted to the specifics of the problem. In addition, the objective function is the aggregate spectral efficiency and cell edge performance. The details of the scenario are summarized in Table 7.

Table 7: Joint Scheduling and Offloading Scenario Aspects

Scenario aspect	Description
Title	LTE-WiFi offloading
Task	T4.2
Network topology	HetNet (macro + outdoor pico)
Radio Access Technology (-ies)	LTE(-A) and WiFi
Nature of small cells	LTE pico/micro/WiFi
Environment	Urban
Context	Both outdoor and indoor areas
Inter-site distance	500 m for macro
Frequency deployment strategy	LTE licensed and WiFi unlicensed bands
Frequency bands	LTE licensed and WiFi unlicensed bands
Density of small cells	Uniform and non-uniform hotspots
Backhaul	X2
Propagation/channel model	A mixture of many models: 3GPP Path loss model for macro and small cells as specified in TR 25.814[i], TR 36.814[ii] Lognormal shadowing with exponential spatial correlation for all links (std and correlation distance will vary). No fast fading for any of the links.
Mobility model	Static users
Traffic model	Full buffer and non-full buffer
Services	Various QoS classes
Number of transmit/receive antennas (for MIMO schemes)	1x1 (downlink)
KPIs, metrics involved	Average throughput per user (-> worst 5th percentile of users -> coverage) Offered area traffic or served area traffic (Mbps/km2, GB/h/km2 -> capacity)
Description of the problem to be solved (target) and proposed method for solution	How to efficiently integrate small cells and WiFi? How to leverage WiFi to improve LTE transmissions
Evaluation method	System level simulations

5.3 Results

Figure 19 plots the convergence behavior of the cross-system learning procedure in terms of the ergodic transmission rate (i.e., average cell throughput). Here, 10 UEs per macrocell sector, and 1.4 MHz bandwidth in the licensed band are considered. In addition, the standard RL algorithm is plotted, in which learning is carried out independently over both licensed and unlicensed bands, without any sort of coordination. Quite remarkably, it is shown that the cross-system learning approach converges within less than 50 iterations, while the standard approach needs several hundreds of iterations to converge. Furthermore, the standard RL procedure exhibits an undesirable oscillating behavior (i.e. ping-pong effects between the licensed and unlicensed band, which can be detrimental in mobility scenarios). Finally, it is worth noting that convergence can be proven using tools from RL.

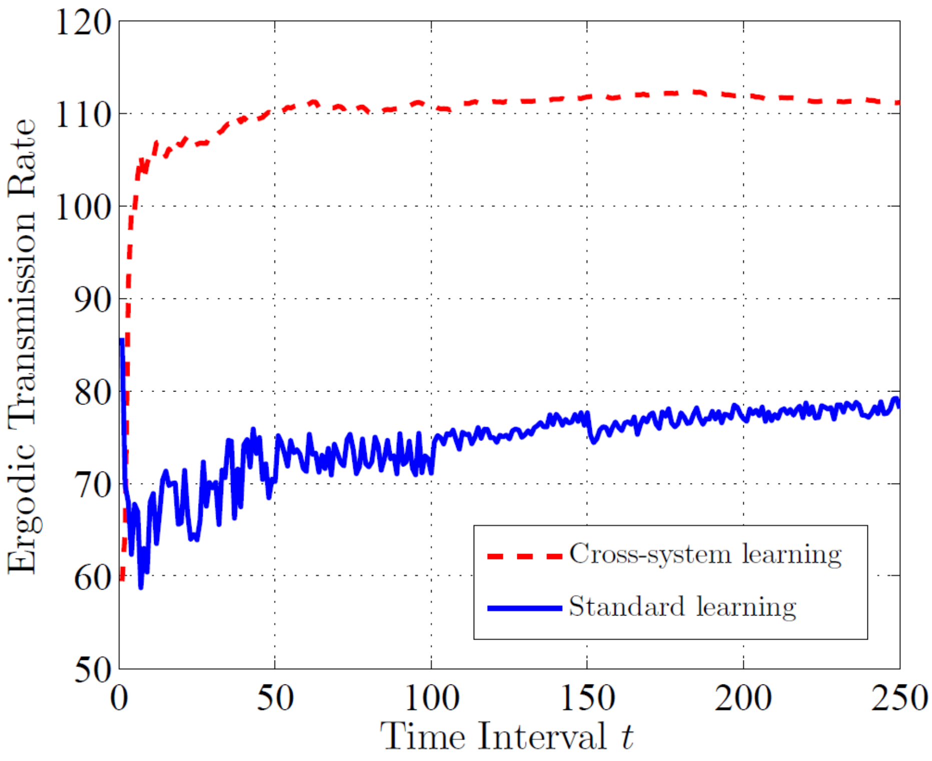


Figure 19 Convergence of the cross-system learning algorithm vs. the standard independent learning, K = 2 SCBSs per macrocell sector

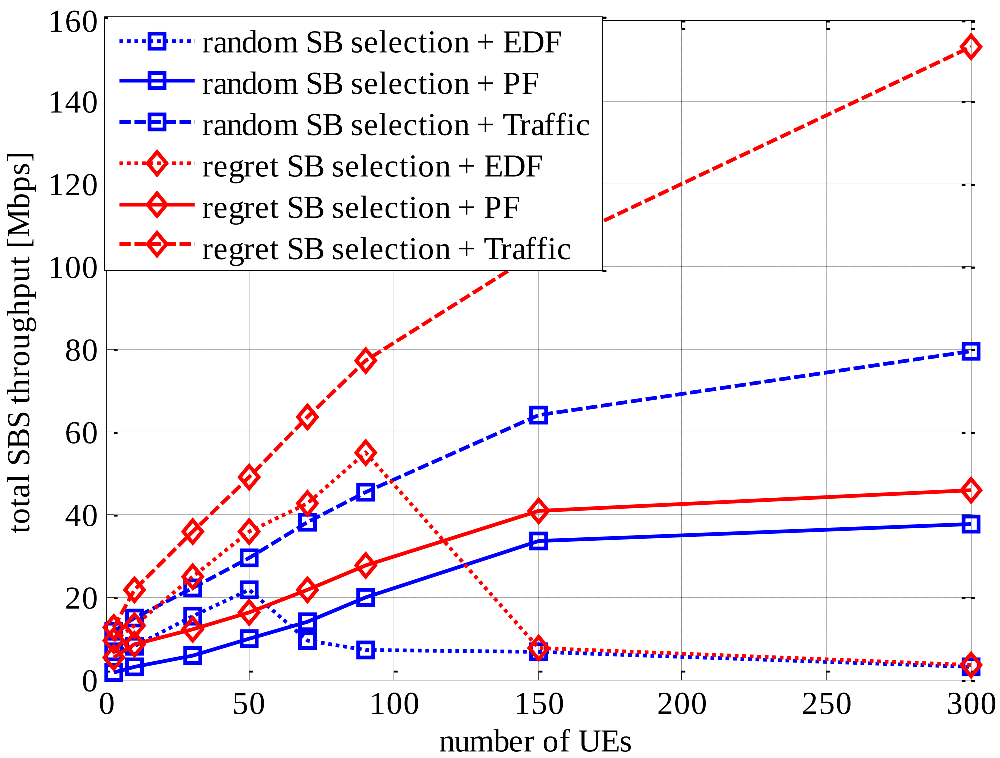


Figure 20 Total cell throughput vs. number of users for different traffic offloading and scheduling strategies, K = 2 SCBSs/macrocell sector.

Figure 20 shows the total cell throughput as a function of the number of UEs in the network, for the earliest deadline first (EDF), proportional fair (PF), and proactive scheduling (PS) strategies, respectively. While the standard PF scheduler cannot cope with the increasing number of UEs, the traffic-aware scheduling approach judiciously steers users' traffic in an intelligent and dynamic manner over both licensed and unlicensed spectrum, with a 160-fold increase for 300 UEs. These significant gains are rooted in the fact that unlike the proactive scheduler, both EDF and PF schedulers fall short of accounting for the heterogeneous traffic and delay-tolerant nature of their users.

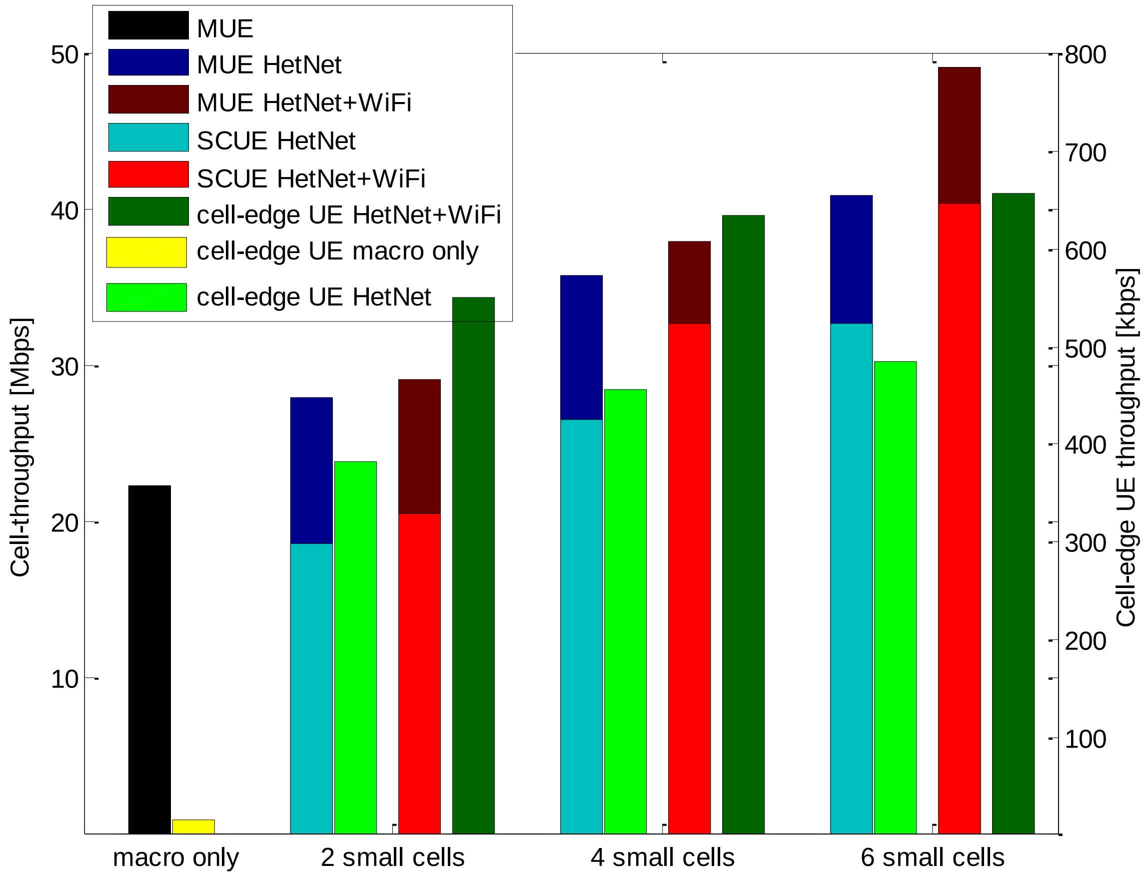


Figure 21 Total cell throughput vs. number of users for different traffic offloading and scheduling strategies, K = 2 SCBSs/macrocell sector.

Figure 21 plots the total cell throughput and cell edge UE throughput for the macro-only, HetNet, and HetNet+WiFi offloading strategies. Some key observations are worth mentioning. While in the macro-only case, cell edge UE devices get rather low throughput gains, adding $K = 2$ small cells is shown to boost users' cell edge throughput in the HetNet offload case. In addition, a 50 percent increase in cell edge UE throughput is obtained with $K = 2$ multimode small cells (HetNet+WiFi). Furthermore, small cell users (SCUEs) benefit from the small cells' multimode capability when deploying $K = 2$ SCBSs, and this gap further increases when adding more small cells ($K = 6$ SCBSs). As a byproduct of this, offloading is shown to improve not only the performance of SCUEs, but also MUEs, for $K = \{2, 4, 6\}$ SCBSs.

6 CONCLUSION

Inter-system radio access offloading, an efficient and cost-effective integration of cellular and WiFi technologies, is an indispensable feature of future dense heterogenous networks. As new macro sites are becoming increasingly difficult and often expensive to deploy, operators will deploy more and more low-power sites within traffic hotspots, densifying future heterogeneous networks.

This deliverable has provided a detailed view of inter-system radio access offloading techniques on innovative concepts for heterogeneous network deployments. In addition to the concept descriptions, detailed performance evaluation results have also been presented.

New concepts are presented within the areas of Inter-LTE traffic offloading via middleware deployment, seamless offloading, capacity aware multi-user offloading, joint offloading and scheduling for dual mode small cells.

In inter-LTE offloading via middleware deployment contribution, a middleware is introduced at the control plane and IP level needed to exchange user context information to intelligently allocate users on the most appropriate cell within the coverage area.

In seamless offloading scenario, the interaction between a 3G network and a WLAN network to make a seamless offload was analyzed and simulated. Based on the results, a high MOS value during the video transmission and better user experience is achieved.

In capacity aware multi-user offloading, multiple-attribute decision making algorithm usage is studied to enhance QoE for a higher percent of users. Based on the results, a balanced load among access networks is achieved with an improved user experience.

In joint offloading and scheduling for dual mode small cells, a cross-system learning approach significantly outperforms a number of benchmark traffic steering policies. Ofloading is shown to improve not only the performance of small cell users, but also macro cell users.

In addition to presenting new concepts within the inter-system radio access offloading (Task 4.2), this deliverable discusses also the performance and deployment strategies of heterogeneous network deployments within various scenarios.

The proposed concepts and innovations on inter-system radio access offloading improve the overall system capacity as well as quality of service. Furthermore, the proposed concepts are also shown to performance improvements in terms of user quality of experience and contribute to cost-effective integration of both WiFi and cellular radio access technologies for operators.

7 LIST OF ABBREVIATIONS, ACRONYMS, AND DEFINITIONS

ACRONYM	DEFINITION
3G	Third Generation cellular system
3GPP	Third Generation Partnership Project
ANDSF	Access Network Discovery and Selection Function
AP	Access Point
BER	Bit Error Rate
BLER	Block Error Rate
BS	Base Station
BW	Bandwidth
CDF	Cumulative Distribution Function
CQI	Channel Quality Indicator
EAP-SIM	Extensible Authentication Protocol – Subscriber Identity Module
EDF	Earliest Deadline First
EPC	Evolved Packet Core
EURANE	Enhanced UMTS Radio Access Network Extension
HetNets	Heterogenous Networks
HIP	Host Identity Protocol
HWN	Heterogeneous Wireless Networks
IP	Internet Protocol
ISD	Inter-Site Distance
ITU-T	International Telecommunication Union – Telecommunication Standardization Sector
KPI	Key Perforamance Indicator
L2	Layer 2
L3	Layer 3
LB	Load Balancing
LTE	3GPP Long Term Evolution
LTE-A	LTE-Advanced
MADM	Multiple Attribute Decision Making

MIH	Media Independent Handover
MIHF	Media Independent Handover Function
MIIS	Media Independent Information Service Server
MIPv4	Mobile-IP version 4
MIPv6	Mobile-IP version 6
MNO	Mobile Network Operator
MOS	Mean Opinion Score
MUE	Macro UE
NIST	National Institute of Standards and Technology
PESQ	Perceptual Evaluation of Speech Quality
PF	Proportional Fair
PS	Proactive Scheduling
PSNR	Peak Signal to Noise Ratio
PU	Public
QoE	Quality of Experience
QoS	Quality of Service
RAT	Radio Access Technology
RB	Resource Block
RL	Reinforcement Learning
RSS	Received Signal Strength
SCBS	Small Cell Base Station
SHARING	Self-Organized Heterogeneous Advanced Radio Networks Generation
SIM	Subscriber Identity Module
SINR	Signal to Interference and Noise Ratio
SIP	Session Initiation Protocol
SNR	Signal to Noise Ratio
TOPSIS	Technique for Order Preference by Similarity to Ideal Solution
UE	User Equipment
UMTS	Universal Mobile Telecommunications System

WEP	Wired Equivalent Privacy
WiFi	Wireless Fidelity
WiMAX	Worldwide Interoperability for Microwave Access
WLAN	Wireless Local Area Network
WP	Work Package
WPA	WiFi Protected Access
WPA2	WiFi Protected Access II

8 References

[Bala10] A. Balasubramanian, R. Mahajan, and A. Venkataramani, " Augmenting mobile 3G using WiFi" In Proceedings of the 8th international conference on Mobile systems, applications, and services, pages 209-222. ACM, 2010

[Bennis13] M. Bennis et al., "When Cellular Meets WiFi in Wireless Small Cell Networks," IEEE Communication Magazine, 51, 6, 44-50, (2013).

[Cis14] Cisco, "Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2013-2018", February 2014.

[Dutta06] Dutta, A., et al., "Seamless Handover across Heterogeneous Networks - An IEEE802.21 Centric Approach" . In: Proceedings of IEEE WPMC (2006)

[Eri13] Ericsson, "Ericsson Mobility Report – On the Pulse of the Networked Society", November 2013.

[Fadel12] M. Fadel, A.S. Ibrahim, H. Elgebaly, "QoS-Aware Multi-RAT resource allocation with minimum transmit power in multiuser OFDM system," Globecom Workshops (GC Wkshps), 2012 IEEE , vol., no., pp.670,675, 3-7 Dec. 2012

[Gang05] X. Gang, G. Min, H. E. Mellor, K. Al-Begain, L. Guan, "An adaptive QoS framework for integrated cellular and WLAN networks", Computer Networks, Volume 47, Issue 2, 4 February 2005, Pages 167-183, ISSN 1389-1286, 10.1016/j.comnet.2004.07.003.

[Hadiji12] Hadiji, F.; Zarai, F.; Kamoun, A., "Architecture of mobile node in heterogeneous networks," Communications and Information Technology (ICCIT), 2012 International Conference on , vol., no., pp.260,264, 26-28 June 2012.

[IEEE21] IEEE Standards Working Group, http://www.ieee802.org/21/, IEEE 802.21 Media Independent Handover Working Group.

[Koun07] G. Koundourakis, D. Axiotis, M. Theologou, "Network-based access selection in composite radio environments" IEEE Wireless Communications and Networking Conference, 2007. WCNC'2007, 11–15 March 2007, pp. 3877– 3883

[Liu06] X. Liu, V.O.K. Li, P. Zhang, "Joint radio resource management through vertical handoffs in 4G networks" in: IEEE Global Telecommunications Conference, 2006. GLOBECOM'06, November 2006, pp. 1–5.

[Logot12] M. Logothetis, K. Tsagkaris, P. Demestichas, "Application and mobility aware integration of opportunistic networks with wireless infrastructures" Computers & Electrical Engineering, Available online 24 August 2012, ISSN 0045-7906, 10.1016/j.compeleceng.2012.07.014

[Mark10] Z. Markovic "Modification of TOPSIS method for solving of multi-criteria tasks". In: YugoslavJournal of Operations Research. 20, p.117-143. (2010)

[Nacef08] A. Ben Nacef, N. Montavont, "A generic end-host mechanism for path selection and flow distribution" in: IEEE 19th International Symposium on Personal, Indoor and Mobile Radio Communications PIMRC, 2008, pp. 1–5.

[Niyato06] D. Niyato, E. Hossain "A Cooperative game framework for bandwidth allocation in 4G heterogeneous wireless networks," in: IEEE International Conference on Communications, 2006. ICC'06, vol. 9, June 2006, pp. 4357–4362.

[Ormond05] O. Ormond, P. Perry, J. Murphy, "Network selection decision in wireless heterogeneous networks" in Proc. of IEEE 16th International Symposium on Personal, Indoor and Mobile Radio Communications, 2005. PIMRC 2005, vol. 4, 11–14 September 2005, pp. 2680–2684.

[Piamrat11] K. Piamrat, A. Ksentini, J. Bonnin, C. Viho, "Radio resource management in emerging heterogeneous wireless networks", Computer Communications, Volume 34, Issue 9, 15 June 2011, Pages 1066-1076, ISSN 0140-3664, 10.1016/j.comcom.2010.02.015.

[Rec01] Rec, I. T. U. T. (2001). P. 862 Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs.

[Rec03] Rec, I. T. U. T. (2003). G. 107. The E-model, a computational model for use in transmission planning.

[ShD41] SHARING Deliverable, “D4.1 New opportunities, challenges and innovative concept candidates for SON/Heterogeneous Networks”, http://www-sharing.cea.fr, June 2014

[Song05] Q. Song, A. Jamalipour, " A network selection mechanism for next generation networks”, IEEE International Conference on Communications, 2005. ICC’2005, vol. 2, 16–20 May 2005, pp. 1418–1422.

[Stevens17] E. Stevens-Navarro and V. W. S. Wong "Comparison between vertical handoff decision algorithms for heterogeneous wireless networks”, Vehicular Technology Conference, VTC 2006-Spring. IEEE 63rd, vol. 2, Melbourne, Vic., May 2006, pp. 947-951

[Taha04] A. Taha, H. Hassanein, H. Mouftah, "On robust allocation policies in wireless heterogeneous networks," in: First International Conference on Quality of Service in Heterogeneous Wired/Wireless Networks, 2004. QSHINE 2004, 18–20 October 2004, pp. 198–205.

[Wilson05] A. Wilson, A. Lenaghan, R. Malyan, "Optimising wireless access network selection to maintain QoS in heterogeneous wireless environments Wireless Personal Multimedia Communications 2005. WPMC’05, 18–22 September 2005.

[Yang05] X. Yang, J. Bigham, L. Cuthbert, "Resource management for service providers in heterogeneous wireless networks” in: IEEE Wireless Communications and Networking Conference 2005, vol. 3, 13–17 March 2005, pp. 1305–1310.