LOAD-AWARE RADIO ACCESS SELECTION IN FUTURE GENERATION SATELLITE-TERRESTRIAL WIRELESS NETWORKS

M. Ali, P Pillai and Y.F.Hu

School of Engineering Design and Technology
University of Bradford, UK

`{m.ali70, p.pillai, y.f.hu}@bradford.ac.uk`

ABSTRACT

In the telecommunication networks the introduction of Next Generation Wireless Networks (NGWN) has been described as the most significant change in wireless communication. The convergence of different access networks in NGWN allows generalized mobility, consistency and ubiquitous provision of services to mobile users. The general target of NGWN is to transport different types of information like voice, data, and other media like video in packets form like IP. The NGWNs offer significant savings in costs to the operators along with new and interesting services to the consumers. Major challenges in NGWN are efficient resource utilization, maintaining service quality, reliability and the security. This paper proposes a solution for seamless load aware Radio Access Technology (RAT) selection based on interworking of different RATs in NGWN. In this paper novel load balancing algorithms have been proposed which have been simulated on the target network architecture for TCP data services. The IEEE 802.21 Media Independent Handover (MIH) is utilized in load balancing specifically for mobility management, which enable low handover latency by reducing the target network detection time. The proposed method considers the network type, signal strength, data rate and network load as primary decision parameters for RAT selection process and consists of two different algorithms, one located in the mobile terminal and the other at the network side. The network architecture, the proposed load balancing framework and RAT selection algorithms were simulated using NS2. Different attributes like load distribution in the wireless networks and average throughput to evaluate the effects of load balancing in considered scenarios.

KEYWORDS

NGWN, Load balancing, radio resource management, heterogeneous wireless networks, load balancing in wireless networks, vertical handovers, and load balancing in satellite-terrestrial wireless networks.

1. INTRODUCTION

Modern mobile devices like cell phones, PDA's, Tablet PCs already support multiple wireless technologies like UMTS, WLAN and Bluetooth and in the very near future would also support satellite and WiMax with multiple interfaces provision. While most of these devices are able to scan the different available networks the user would manually select which network he or she may want to use. It is envisaged that in the NGWN these devices may be able to apply some complex Radio Access Technology (RAT) selection techniques to find the most suitable network. Such a RAT selection technique may need to consider various parameters like received signal strengths, errors rates, costs, user preferences, QoS requirements, etc. Such a

35

RAT selection technique would not only play an important part when a user switches on his or her mobile device but also when the user moves around. While most of the current day mobile networks already support seamless handovers, these are restricted to handovers within the same technology, i.e. horizontal handovers. It is envisaged that to efficiently use the network services the future mobile devices shall also support handovers across different radio access technologies. This process of switching mobile devices connectivity from one technology to another type of technology is called vertical handover. The joint call admission control (JCAC) algorithm for next generation heterogeneous wireless networks is envisioned as user-centric. User centricity implies that user's preferences are considered in decision making for RAT selection. However user-centric JCAC algorithms often lead to highly unbalanced networks load, which cause congestion on overloaded network and eventually increase the call blocking and call dropping probabilities. The unbalanced load situation in co-located networks also causes the poor radio resource utilization as some networks remain under loaded and some get over loaded. The load balancing strategies are required to efficiently utilize the available radio resources and avoid the unwanted congestion situations due to overloaded wireless networks. This paper presents a novel NGWN RAT selection technique which uniformly distributes the network load between co-located heterogeneous wireless networks. It utilizes IEEE 802.21 Media independent Handover (MIH)[1] to seamlessly handover mobile users between heterogeneous wireless networks for load balancing purpose. The advantage of this approach is that it minimizes the call blocking and dropping probabilities, number of packet drop/lost and delays during the handover process and enhances the network utilization by continuously balancing the load in co-located networks. The proposed load balancing approach monitors and controls the network load from both side (mobile node and network side), and addresses the most important problem in NGWN which is efficient radio resource utilization. The rest of the paper is organized as follows; section 2 describes the literature review of load balancing strategies in detail, section 3 briefly describes the proposed load balancing framework. The protocol stack and detailed handover procedures are described insection 4. Finally simulation scenarios and results are discussed in section 5, which is followed by the conclusion.

2. LOAD BALANCING STRATEGIES

Usually more than one wireless networks may provide coverage to any given location in an urban area. For example, when working in an office building, the mobile device of a user may be in the coverage of a UMTS mobile network and a WLAN office network. A user may manually configure to use the UMTs network for voice services and the WLAN access for data services. In such overlapping coverage areas of different wireless networks such as satellite networks and terrestrial networks like WiMax, UMTS and WLAN; a RAT selection technique is required to find the most suitable network based on received signal strengths, bit errors rates, costs, user preferences, QoS requirements, and most importantly the load of networks. A detailed description of load balancing mechanism in heterogeneous wireless networks is described in the following subsections.

2.1 Load balancing mechanism

The load balancing mechanism can generally be divided into two main parts in heterogeneous wireless networks [2]; the load balancing algorithm and the network architecture. The later part which is network architecture is the base for the efficient load balancing and good network architecture can improve the efficiency of load balancing mechanism. From the control mode perspective the load balancing mechanism can be categorized as distributed, semi-centralized

semi-distributed and centralized load balancing mechanism [3, 4]. Both pure centralized and distributed approaches have issues as distributed approach in the heterogeneous wireless network architecture will have a huge signalling overhead and the centralized approach will have low reliability [5].

2.2 Load balancing architectures

In [4] the authors proposed a semi-distributed and semi-centralized architecture (SCSDA), which is used in such a way that BS's exchange load information with other neighbouring BS's. This architecture in theory is able to reduce the overhead of control signalling but it was not proved in the paper by the authors using simulation or analytical model. A hybrid wireless network architecture design presented in [6] and whereas network architecture based on multiple mobile routers to support seamless mobility across future heterogeneous wireless network is presented in [7]. Both [6] and [7] verified the reduced overhead by NS2 simulations however the model was not derived in these approaches. Theoretical route overhead is presented in the [8] by counting the number of control messages generated in the network entities such as BS or AP for route maintenance. In [9] the communication overhead of the presented mechanism was calculated and to minimize this communication overhead an algorithm was presented. The presented algorithm was proved effective with the help of simulation results. The general heterogeneous wireless network was considered in the [10], where two basic network entities the mobile node and the AP are considered. This approach formulated the overhead for the discovery of AP by dividing it into RREQ messages and hello messages and proved the effectiveness of their proposed method with help of simulation results. The mechanism presented in the [11] proposed a hierarchical distributed architecture with three levels of hierarchy in mobility management. The three levels of mobility are as follows: i) the end terminal changes its point of attachment but remains connected to the same radio access network. ii) The end terminal remains associated to the same operator but changes its radio access network. iii) The end terminal changes its operator network. In this paper the authors also estimated the signalling cost during the QoS negotiation for handover process. In [12] a hierarchical semi-centralized architecture is presented which considered heterogeneous wireless networks including WiMax, Wi-Fi and UMTS to share network load. This approach introduced new entities like resource allocator, information servers and resource statistics and named them collectively as Resource Management Unit (RMU). The authors have also provided list of signallings between the newly introduced entities in the network architecture. The disadvantage of this approach is that it is not standardized approach for integration of heterogeneous wireless networks and requires upgrades in wireless networks architecture as well as in protocol stacks. On the other hand the proposed approach in [13, 14, 15, 16] have also adapted semi-centralized architecture but utilized the enhanced MIH [1] which is a standardized and more efficient. In [17] the authors adapted the loosely coupled architecture for the cellular/WLAN integration [18] for the load balancing purpose between heterogeneous wireless networks. The authors adopted the two phase control strategies in the load balancing policies. The dynamic vertical handover during the traffic serving phase is used to make the performance variance smooth and call admission is used to provision static QoS guarantee during the admission phase. The effectiveness of proposed strategy in this approach was proved with the statistical comparison of results with other similarly presented references. In [19] the authors presented a dynamic load balancing architecture for the load balancing in heterogeneous wireless access networks which considers WiMax, 3GPP Long Term Evolution (LTE) and Wi-Fi. This approach adapted the central architecture for load balancing and introduced new entities in the wireless networks architecture, like Community resource manager (CRM), Local resource manager (LRM),

community access point (CAP) and spectrum manager (SM). The approaches presented in [20, 21, 22] discussed the inter RAT load balancing algorithms and [23] presented the usage of radio enabler in IEEE-P1900.4. The discovery of RATs is assisted from the information provided by the radio enabler. In [27] load balancing approach has been presented which targets the proxy mobile ipv6 (PMIPV6) domain using MIH for heterogeneous networks. A comparison has been made between the scenario performing load balancing in extended PMIPV6 for handover signalling and the scenario using MIH signalling for load balancing. It was shown in the results that use of load balancing improves the efficiency whereas, MIH based load balancing improves data rate as compared to extended MIPV6 based load balancing. This disadvantage in this approach is when considering load-aware RAT selection; it is specifically designed for a MIPV6 architecture using Local Mobility Agent (LMA) and a new entity called Mobile Access Gateway (MAG) in the network. The proposed architecture in this paper for load balancing is semi-centralized and semi-distributed as the load balancing is performed at both ends such as mobile node and the network entity like BS, AP, RAN and RNC.

2.3 Load balancing algorithms

The load balancing approaches presented in [24] and [25] have considered load balancing in homogenous network targeting WLAN. The approach in [24] considers the received signal strength indicator (RSSI) value to distribute the load between different access points (AP's) which have overlapping coverage areas. This approach uses the two values in balancing the load which are RSSI between mobile station (MS) and AP and the average RSSI value of all the MS's currently connected with AP. The method given in [26] considers both RSSI and the number of MS associated with AP which makes it much effective for load balancing. The technique used in [26] presented a solution for load balancing in homogeneous wireless networks, by utilizing genetic algorithm. As the genetic algorithm's convergence directly proportional to the size of population (mobile nodes and APs) therefore this approach is effective for WLAN networks and not for the heterogeneous wireless environment where population size is comparatively large due to large coverage areas. All approaches given in [24, 25, 26] were designed to enhance the performance for homogeneous network environment particularly for WLAN.In [27] load balancing approach has been presented which targets the proxy mobile ipv6 (PMIPV6) domain using MIH for heterogeneous networks. A comparison has been made between the scenario performing load balancing in extended PMIPV6 for handover signalling and the scenario using MIH signalling for load balancing. It was shown in the results that use of load balancing improves the efficiency whereas, MIH based load balancing improves data rate as compared to extended MIPV6 based load balancing. This disadvantage in this approach is when considering load-aware RAT selection; it is specifically designed for a MIPV6 architecture using Local Mobility Agent (LMA) and a new entity called Mobile Access Gateway (MAG) in the network. In [28] a general set of algorithms have been proposed which considers battery power of mobile users, received signal strength and load on available points of attachments in handover process to balance the load in co-located networks overlapping their coverage areas. In this approach load balancing is done only at network side without any interaction with the mobile node. On the other hand our proposed approach considers both; mobile nodes and network entities such as AP, BS and satellite ground station for load balancing thereby resulting in more efficient load balancing across the neighbouring networks. In [29] a detailed algorithm has been presented for network selection in heterogeneous wireless networks. The algorithm presented in [29] has been divided into two parts, one runs at mobile terminals and other part of algorithm runs at network entity such as basestation (BS) or access point (AP). This approach considers received signal-strength, battery

power, speed, and location of mobile user but does not considers MIH which could have improved the handover process while moving the mobile nodes between different networks. In [30] a next generation networks (NGN) based approach has been presented in which hierarchical joint call admission control algorithm is extended to send newly added load reports from hierarchical call admission control (HCAC) entity to vertical call admission control entity (VCAC). The main goals of proposed approach in [30] are simplicity and scalability, however this approach performs balancing of load periodically and therefore may not performs very efficiently with abrupt load changes in different sub networks in the hierarchy. In [31] a Markov chain based model for load balancing and QoS based CAC has been presented and comparisons have been made between the results of load balancing based CAC and QoS based CAC algorithms. The load balancing approach presented in this research report is more efficient than load based CAC approach presented in [31] as our approach uses MIH to minimize the handover delays when moving the mobile nodes for load balancing purpose and tends to uniformly distribute the load among available heterogeneous wireless networks.

3. PROPOSED LOAD-AWARE RAT SELECTION FRAMEWORK

3.1 Proposed Network Architecture

Figure 1 presents the target network architecture which is considered in this paper. It shows an MIH enabled multi interface mobile device which can use any of the three available wireless networks supported by its interfaces.
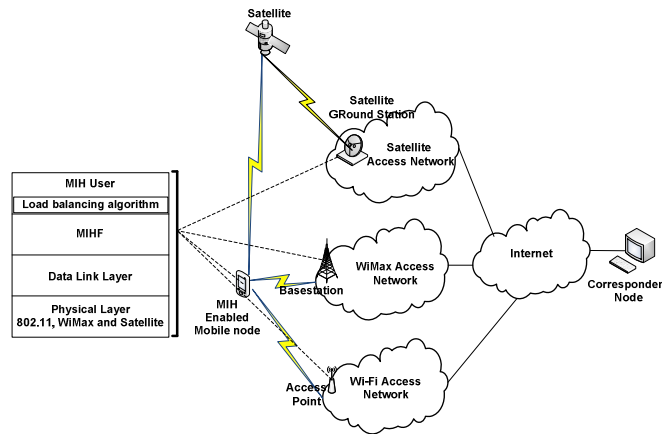

Figure 1: Target network architecture

The access network of each technology such as Satellite, WiMax and Wi-Fi is connected to internet. There is also a correspondent node located behind the internet as shown in the Figure 1. The mobile node can communicate with the correspondent node over the internet using any available network which is supported by its interfaces. The mobile node handovers to different available networks while moving from coverage area of one network to another and during this mobility it can maintain the communication with correspondent node. The load balancing algorithms are located at the MIH user in MIH reference model as represented by the Figure 1. MIH user is selected for the load balancing process origin as MIH user is the central control point for triggering and handling MIH signalling as described in [1]. In mobile user the load balancing algorithm shown in Figure 2 is adopted and in Satellite Ground Station/BS/AP the load balancing algorithm for the network entity shown in Figure 3 is adopted.

3.2 Proposed Load balancing algorithm

This section describes the proposed load-aware RAT selection algorithm. The proposed algorithm considers the network type, signal strength, data rate and network load as primary decision parameters for RAT selection process and tries to maintain the load equilibrium on all networks which have common or overlapped coverage areas. It is assumed that all considered networks and mobile nodes support the IEEE 802.21 MIH. The IEEE 802.21 MIH standard has been brought into use for seamless vertical handover operations of mobile nodes between the co-located wireless networks. The proposed approach has taken advantage of MIH media independent information service (MIIS) specifically for exchange of network load information besides exchanging other network related information like link type, link data rate, link capability, offered security and QoS and cost[1]. The proposed RAT selection framework consists of two segments of load aware algorithm, one running on mobile device and other running on network entity like Satellite Ground Station, WiMax BS or WLAN AP. The flow chart shown in Figure 2 represents the proposed algorithm's which runs at mobile device. At the mobile device, the proposed technique first makes a list of available network IDs which are visible to mobile device such that received signal strength from those networks is higher than the minimum threshold. In next step load value of each network in the list is obtained from MIIS and compared.
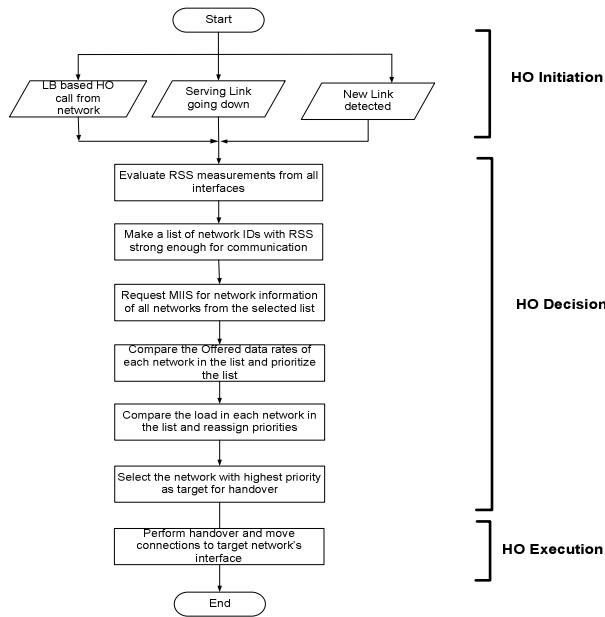

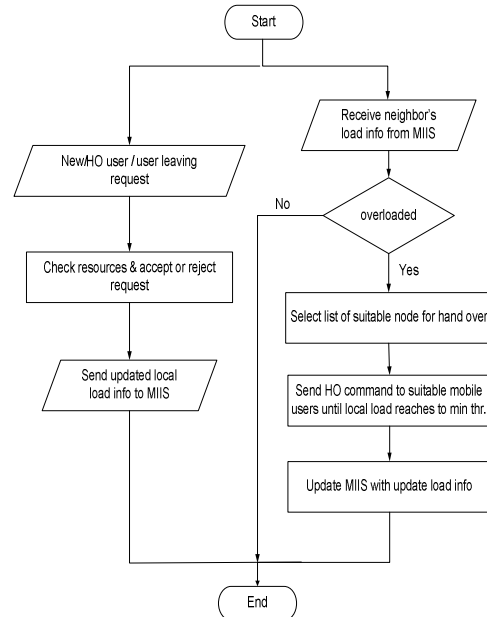Figure 2: LB algo. in the mobile device   Figure 3: LB algo. at the Network side

Then in following step it compares the data rate offered by each network in the list. The most preferred network from the list is the one with lowest load and highest offered data rate. The second algorithm shown in Figure 3 runs in network side. In the network entity like BS or AP the load balancing algorithm continuously keeps on updating the MIIS about its current load status and receives load information of its neighbouring networks. This updating process runs on every time when a new connection starts or ends in the network. The most loaded network

entity start moving out the suitable mobile users to appropriate networks, if the load variation is gone higher than threshold of 50% free resources margin, such that the percentage of free resources in one network is greater than or equal to the double of available resources percentage at any other network. Load balancing algorithm keeps on migrating out the suitable mobile nodes from over loaded network to the least loaded networks until the load in over loaded network becomes equal to or lesser than the average load in all the neighbouring networks of overloaded network. The load balancing is performed by the handover procedures in both mobile and network side. In mobile nodes it is supported by the mobile node initiated handovers and on network entity it is supported by the network initiated handovers of the selected mobile nodes in the network.

4. PROTOCOL STACK AND HANDOVER PROCEDURES

4.1 Protocol Stack

The Figure 1 represents the location of load balancing algorithm in the MIH reference model. The key phenomenon in the MIH reference model is the introduction of Media Independent Handover Function (MIHF) between layer 2 and layer 3 of the Open System Interconnection (OSI) layer model. The MIHF can be defined as the newly created entity within the protocol stack of network elements like mobile node (MN) or the network entity such as RNC, NCC or RAN, BS and AP.
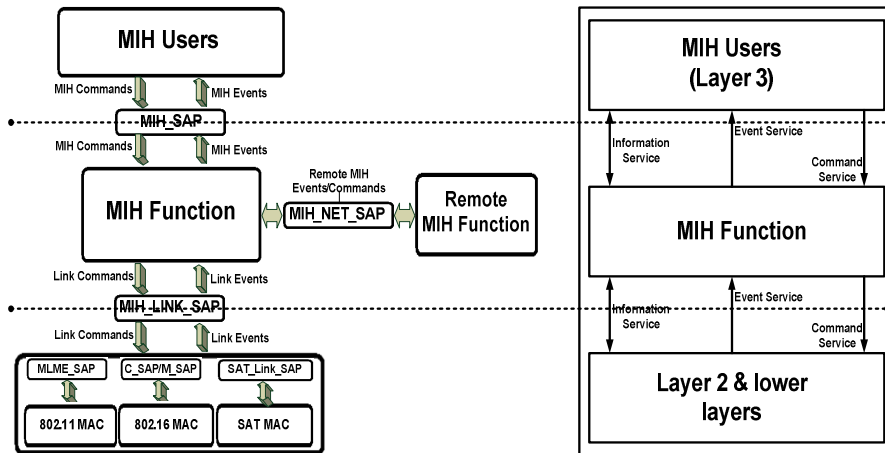

Figure 4: MIH reference model in OSI layers

The MIHF is composed of a set of handover enabling functions which provide service continuity while a MN traverses between heterogeneous wireless access link layer technologies. The Figure 4 shows an analogy of MIHF and the information flow within the MIH reference model with respect to the OSI layer. It shows that in the protocol stack the MIHF is situated between layer 2 and layer 3 and uses the Service Access Points (SAPS) to interface the higher and lower layers with MIHF. Each SAP is composed of a service primitives set which specifies the interactions between the provider and the service user. Each primitive in the SAP contains the list of allowed parameters. These parameters are defined using the abstract data types. The MIHF uses two interfaces MIH_SAP and MIH_Link_SAP to communicate with higher layer 3 and the lower layer 2 (MAC layer). The media independent MIH_SAP provides interface between MIHF and the layer 3, whereas the media dependent MIH_Link_SAP provides the

interface between MAC layer (layer 2) and MIHF. This SAP defines the abstract media dependent interface between MIHF and different link layers technologies. In [1] suggestions for the amendments have been given for different link layer technology specific SAPs which are based on the definition of this particular SAP. The link layer specific SAPs for different radio access technologies such as IEEE 802.11, IEEE 802.16, UMTS and Satellites inherit the generic primitives from MIH_Link_SAP. The Figure 4 also shows the link specific SAPs for each of the access technologies considered in this research report. The MLME_SAP provides interface between MIHF and the MAC layer of IEEE802.11, C_SAP and M_SAP provide interface between MIHF and IEEE 802.16 MAC layer, MIH_3GLink_SAP provides the interface between MIHF and UMTS MAC layer and the last newly proposed SAT_Link_SAP provides the interface between satellite MAC and the MIHF.

4.2 Detailed handover procedures

The handover scenarios considered in this section cover the vertical handover procedures between Satellite networks, WiMax and Wi-Fi networks in such a way that it covers circumstances which trigger handover when mobile node enters and leaves these networks. It is assumed in the following handover procedures that the mobile node supports the multi-interfaces and can utilize and network such as satellite, WiMax and Wi-Fi upon detection. It is also assumed that the mobile node establishes and maintains a connection with the remote source which is located beyond the core network [13,16].
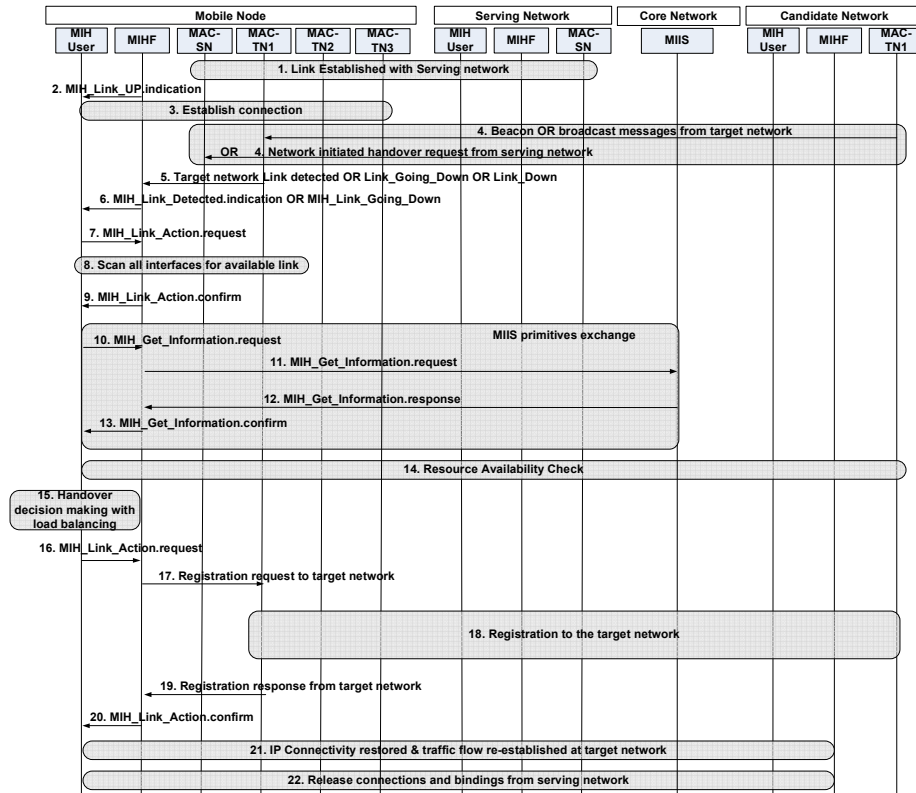

Figure 5: General scenario forHandover (Mobile initiated)

The Figure 5 represents the general handover scenario message sequence chart in which a mobile node handovers from a serving network to a target network after establishing a connection with the serving network. The terms used in the Figure 5 are MAC-SN which represents the MAC layer for serving networks MAC-TN1, MAC-TN2 and MAC-TN3 stands for MAC layer of target network 1, target network 2 and target network 3. The step 1 in the Figure 5 shows that link is established with the serving network or the mobile node has registered with the serving network. The step 2 informs that MIH user in the mobile node about the successful registration of the serving network. The successful connection establishment is shown in the step 3. The step 4 shows the cause of handover or the handover trigger which can either be the detection of new network or the network initiated handover call from the serving network. In step 5 and 6 the MIH user is informed about the handover trigger which is received in step 4. The sequence of message from step 7 to the step 9 represents the procedure of scanning all the interface of mobile node for the available networks. In the messages from step 10 to step 13 the MIH user at mobile node extracts the information about the detected networks which were detected in scanning process. The step 14 shows the resource availability process in the detected networks [1].

For load balancing from network side, the handover procedure is triggered from the network side which is depicted in the message sequence chart shown as Figure 6. This message sequence chart is representing the general diagram for network initiated handover for load balancing purpose. The step 1 in the Figure 6 represents the situation where the network updates MIIS about its load status after change in available resource due to an active connection has been released upon completion or handover, or a new connection have been established. The network also requests for the load information of the neighbouring networks from MIIS at the same time, which is depicted in the sequence of messages until step 5. In step 6 the serving network analyse the load information which is received from the MIIS and decides whether load balancing HOs should be executed or not. Once the decision is made to trigger the load balancing based HOs from network side, the serving network collects the available networks from the suitable nodes which are shown in sequence of steps from step 7 to step 12. The network then instructs the selected mobile nodes to initiate HO process to their candidate networks having lower load. The sequence of steps from step 13 to step 17 represents the network indicating mobile node to handover. The step 18 represents the final operation where mobile node releases bindings with serving network after registering with target candidate networks and establishing traffic flows.
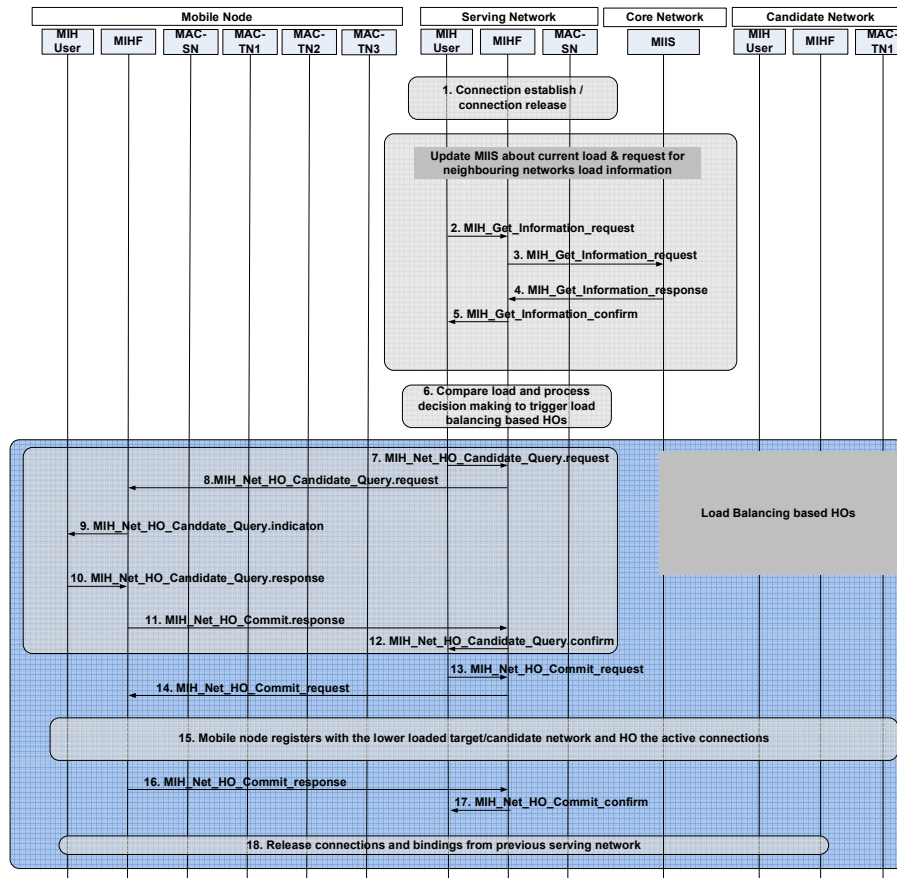
Figure 6: General scenario for Handover (Network initiated)

The target network selection decision is made in the step 15 by utilizing the load aware handover algorithm. The mobile node registers itself with the selected network by using its corresponding interface in the sequence of steps from step 16 to the step 20. Step 21 shows the connections handover from the serving network to the selected target network. All the bindings with the old serving network are released in the step 22.

Figure 7 shows the SAP primitives used in the handover procedure from WiMax to Wi-Fi. As represented by Figure 7 the 802.11 MAC layer in the mobile node, after detecting and registering with Wi-Fi network, it sends MLME-LinkUp.indication message to the MIHF. In step 3, MIHF sends MIH_Link_UP.indication to MIH User. A set of messages from step 4 to 7 acquire the neighbouring networks information. Step 8 checks for the required resources in Wi-Fi for handover. The MIH User decides whether to perform handover or not in step 9. Steps 10 to 12 show the handing over of the connections to the Wi-Fi network. Finally, step 13 and step 14 make sure that traffic flow has been re-establish between mobile node and source and then release bindings with WiMax network. Figure 8 shows the handover procedure when mobile user moves away from the Wi-Fi coverage area and enters a WiMax coverage area. The first step in Figure 8 is the message MLME_MREPORT.indication from MAC Wi-Fi to MIHF. This is the periodic message which carries parameters of link. In step 2 the MIH User is being updated with link parameters report. Step 3 shows the message link-Going_down from Wi-Fi MAC to MIHF, which represents that mobile node, is gradually losing the connectivity with

Wi-Fi. Step 4 informs the MIH User about link going down event. From step 5 to step 8 the messages are used to acquire neighbouring networks information from MIIS. Step 9 shown as bubble represents the process of scanning on all interfaces supported by mobile node. Step 10 and step 11 are for selecting the WiMax network and handover all active connections to WiMax.
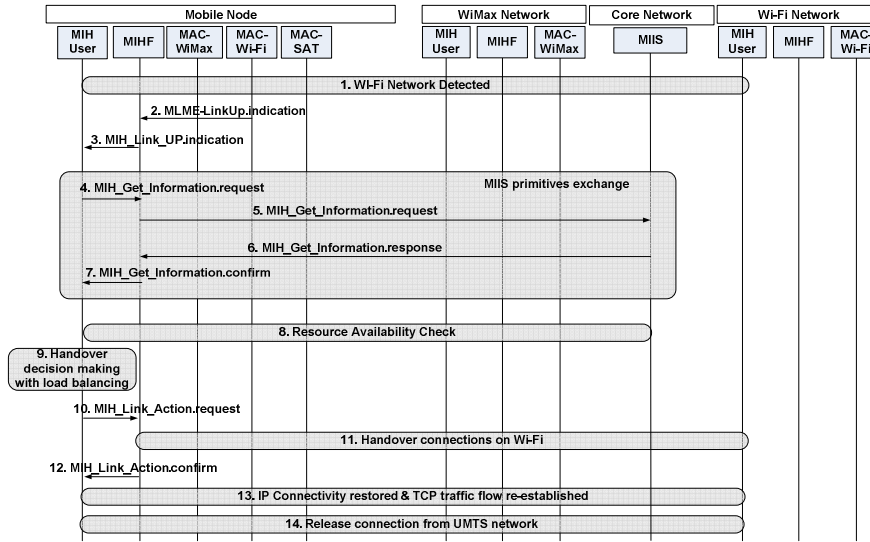

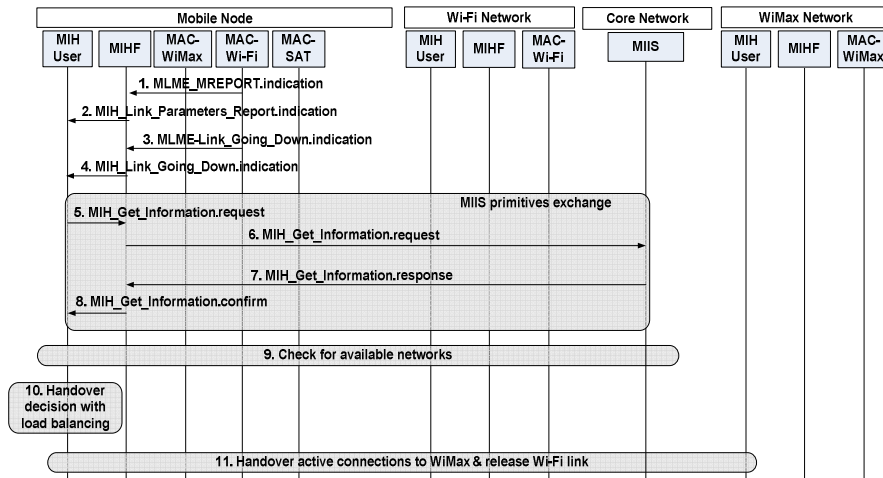Figure 7: Handover from WiMax to Wi-Fi

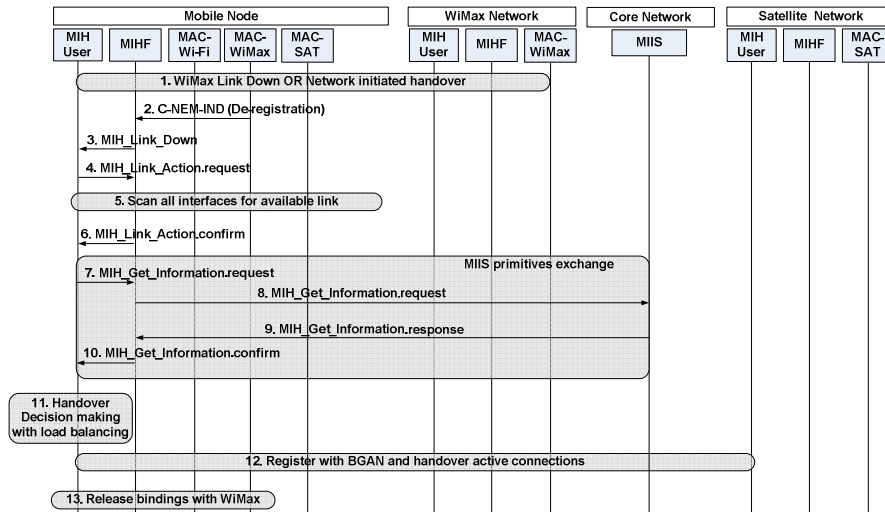

Figure 8: Handover from Wi-Fi to WiMax

Figure 9: Handover from WiMax to Satellite

Figure 9 represents the scenario where mobile node performs the handover operation from WiMax to the satellite network. In the step 1 of the Figure 9 it is assumed that mobile node has lost the connectivity with WiMax network or the WiMax network has sent the De-Registeration message to the mobile node for the load balancing purpose. Step 2 shows that MAC layer of WiMax at mobile node forwarded the De-Registration message to the MIHF which triggers the MIH_Link_Down in step 3 from MIHF to MIH user. The sequence messages from step 4 to the step 6 in Figure 9 represent the process of scanning on all the interface of mobile node and sending scanning reports towards the MIH user. The steps from 7 to 10 are used to gather information from the MIIS about the networks which were detected during the scanning operation at mobile node. The handover decision is made in step 11 and the target network is selected for handover operation. In step 12 the mobile node register itself on network and handover the active connections to the target network. All the bindings with WiMax network are released in step 13. The Figure 10 shows the message sequence chart for the handover process of a mobile node from the satellite network to the WiMax network. It is assumed that the mobile node enters the coverage area of the WiMax network which is underutilized or the mobile node lost connectivity with the currently connected satellite network orthe network initiated the mobile node handover for the load balancing purpose. The step 1 in Figure 10 shows that mobile node lost the connectivity with network. Step 2 represents the message generated from MIHF towards MIH user upon detecting the connectivity lost at MIHF. Step 3 shows the process where MIH user instructs all the interface of mobile node to perform scanning. In step 4 the WiMax interface receives the beacon messages from the WiMax network. Step 5 and 6 forward the received network information of the WiMax network to the MIH user from the MAC layer of WiMax in mobile node via MIHF. From step 7 to the step 10 the MIH user in the mobile node extracts the information of the detected networks including WiMax from the MIIS. Once all the information is gathered at the MIH user the decision for the handover is made which is shown in the step 12. The sequence of steps from step 13 to step 22 represents the registration of mobile node at WiMax network. The bubble at step 23 shows the process of re-establishing the active connection over WiMax. The bindings with old serving network are released at step 24.
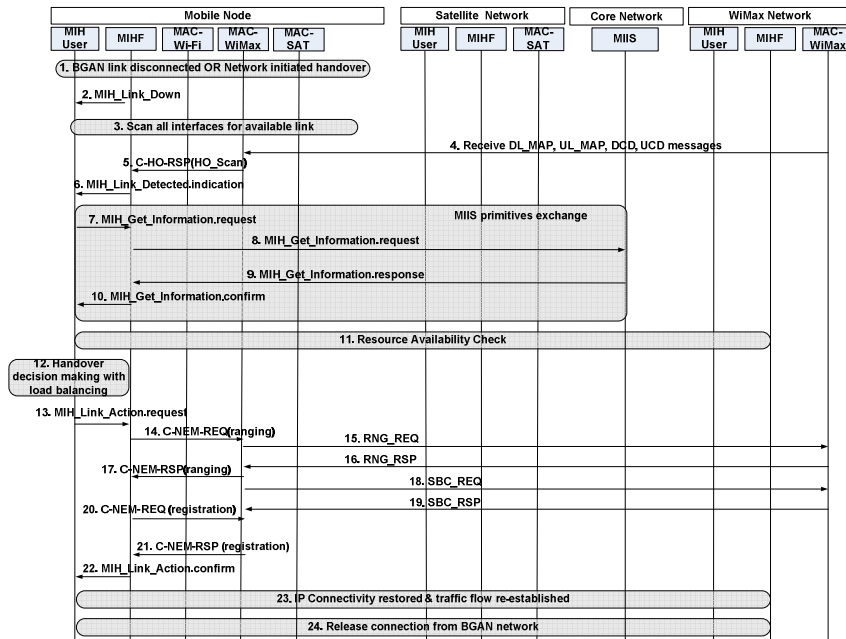
Figure 10: Handover from Satellite to WiMax

5. SIMULATION SCENARIOS&RESULTS

5.1 Simulation Scenarios

Figure 11presents the topology for simulation scenarios considered in this paper. Purpose for considering particular topology for simulation is to observe the effects of load balancing in most ideal scenarios where mobile nodes can see maximum overlapped coverage areas from different wireless networks. Each mobile user maintains a TCP connection with the TCP source shown in Figure 11throughout the simulationsuch that effects of handovers on active connections can be measured.
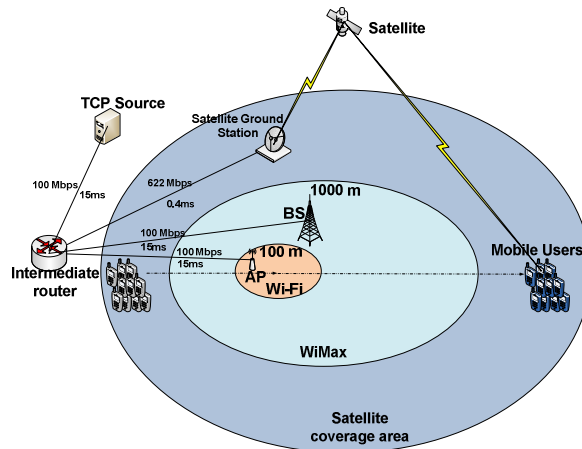

Figure 11: Network topology for simulation

The scenarios considered in this paper consist of a group of mobile users which travel across the coverage areas of all three networks such as Satellite, Wi-Fi and WiMax as shown in Figure 11. In simulation a group of mobile users starts from the Satellite coverage area and move together towards the WiMax coverage area. At time 25 seconds all mobile users enter in WiMax coverage area and at approximately at 170 seconds they leave WiMax coverage area. The Wi-Fi coverage area is overlapped by WiMax therefore at time 96 seconds group of mobile users enters the Wi-Fi coverage area and at approximately 122 seconds all mobile users leave Wi-Fi coverage. Satellite coverage is available to the mobile users throughout the simulation time from time 0 seconds to 200 seconds. All mobile users support multiple interfaces such as Satellite, WiMax and Wi-Fi and use these interface throughout the simulation. Each of these scenarios simulated with load balancing and without load balancing algorithm.

Table 1: Simulation parameters

Parameters	Value
Satellite Coverage	Global
WiMax Coverage	1000m
Wi-Fi Coverage	100m
Speed of mobile nodes	15 m/s
Number of connections per mobile node	1
Connection Type	TCP
Total number of scenarios	4 i.e. {(5nodes,10 nodes)x(LB, NLB)}
Number of mobile users in each scenario	5, 10
Types of technologies supported by each mobile node	Satellite, WiMax and Wi-Fi
Simulation time	200 Sec.

When using the load balancing algorithm in RAT selection; the mobile users are assigned to the network by considering the overall load variation in the co-located networks; therefore it maintains the load equilibrium in the co-located networks. On other hand where load balancing approach is not utilized the mobile users move to the network with the strong signalling strength and high data rate which creates a noticeable amount of load variation in the co-located neighbouring networks. The table 1 represents the simulation parameters for the considered simulation scenarios.

5.2 Simulation Results

5.3 Load balancing performance

All four scenarios discussed in the previous section have been simulated using NS2[32]. Figure 12 and Figure 13represent the load distribution at each network in 5 and 10 nodes scenarios with load balancing and without load balancing. In Figure 12 the blue and the green lines represent the load at Satellite network with and without load balancing. The graph shows that with load balancing Satellite network carries some load throughout the simulation and does not stay idle at all, whereas without load balancing Satellite network carries maximum load at the beginning and end of simulation and stays idle for a long period of time in between; which is poor utilization of radio resource available in the form of Satellite network if the cost is not the factor.Similarly red and cyan lines show the network load for WiMax with and without load balancing respectively. For WiMax the graphs shows that without load balancing all the mobile

nodes move to WiMax when it is available to them. After some time when Wi-Fi becomes available all mobile nodes move to Wi-Fi and WiMax become idle. Upon leaving the coverage area of Wi-Fi all the mobile nodes again join the WiMax network. With load balancing algorithm it can be noticed that WiMax network never stay idle as long as mobile nodes remain in the WiMax coverage area.It shows that with load balancing the WiMax network is utilized better as it never stays idle and not all the mobile nodes connect to it any time. The pink and the yellow lines representing load at Wi-Fi show low load value with load balancing and high load without load balancing. Table 2 shows the comparison of load in different networks for the simulation scenario of 10 nodes with and without load balancing.
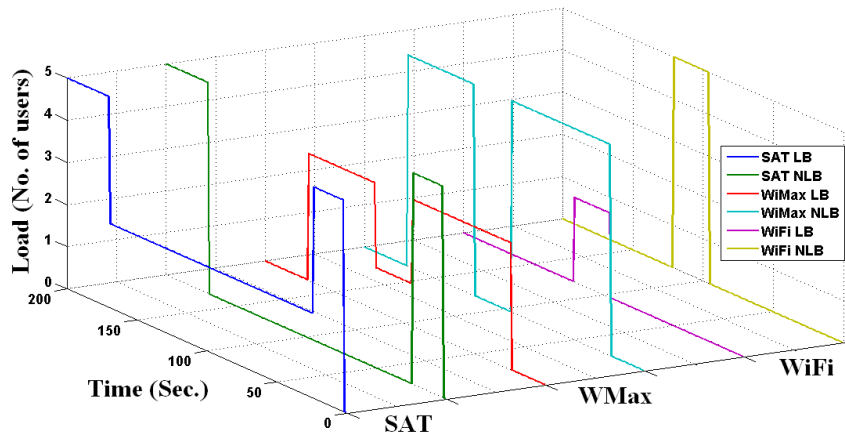

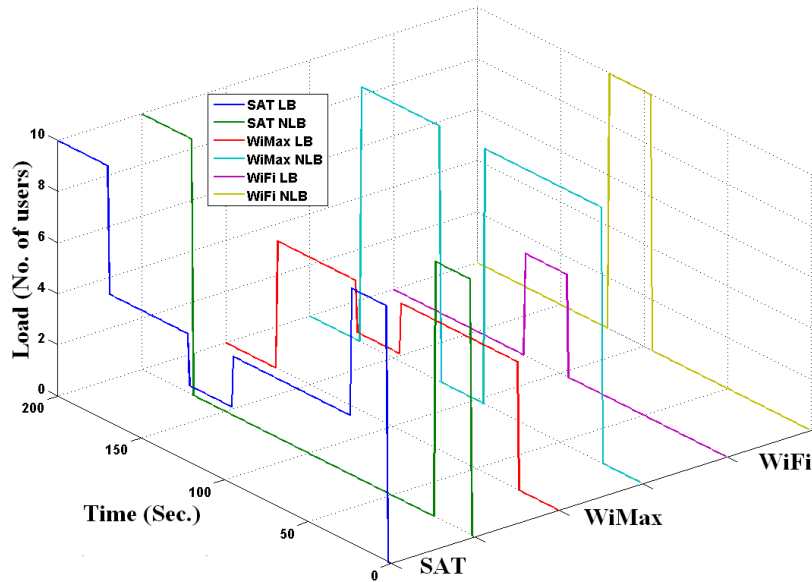Figure 12: Networks load in 5 nodes scenario

Figure 13: Networks load in 10 nodes scenario

Load balancing also minimizes the connection/call blocking and dropping probabilities by sharing the load and maximizing the available networks resources. The term LB in the following Table 2 represents Load Balancing and NLB represents No Load Balancing or without load balancing.

Table 2: Load Table for 10 nodes scenario

Approximate Time	Available Coverage	Sat. Load		WiMax Load		Wi-Fi Load	
		NLB	LB	NLB	LB	NLB	LB
3-25 seconds	Satellite only	10	10	0	0	0	0
25-96 seconds	Satellite& WiMax	0	5	10	5	0	0
96-122	Satellite, WiMax & Wi-Fi	0	3	0	3	10	4
122-170	Satellite& WiMax	0	5	10	5	0	0
170-200	Satellite only	10	10	0	0	0	0

It is shown in the Figure 13 that without load balancing algorithm there have been a lot of load variation such that all the mobile users acquire the best available network, which results in one network overloaded and other nearly idle. On other hand if load balancing algorithm is used all the available networks are equally loaded, which gives the best utilization of available technologies and avoids the situation where one network gets overloaded while other having minimum load. For example in case of 10 nodes scenario without load balancing from time 3 to 25 sec Satellitenetwork has all 10 mobile users at time 25 sec all 10 users handover to WiMax and Satellite load becomes 0. Then at time 96 sec all 10 users handover to Wi-Fi and load of WiMax and Satellite becomes 0. When all mobile users leave the Wi-Fi coverage area at time 122 sec Wi-Fi load becomes 0 and WiMax load becomes 10. On other hand with load balancing the 10 mobile users scenario shows different load distribution. Load at Satellite goes to 10 initially but as WiMax becomes available Satellite load becomes 5 and WiMax load becomes 5such that each available network gets equal load. After some time when Wi-Fi becomes available to the mobile users then load at Satellite and WiMax becomes 3 and load at Wi-Fi becomes 4. On leaving from the Wi-Fi coverage area the mobile users from Wi-Fi network move toSatellite and WiMax such that the load at Satellite becomes 5 and load at WiMax becomes 5.The reason for such variations in the results for the scenarios with and without load balancing is that without load balancing all mobile users move to the network with strong signal strength and high data rate, leaving other available networks in that area with low or no load and over populating the best available network as shown in the list of figures from Figure 12 to Figure 13. On other hand when applying the load balancing algorithm, it controls the unbalanced distribution of mobile nodes to the available networks. The graphs in the figures (from Figure 12 to Figure 13) are showing that load balancing maximizes the chances of availability of any network and minimizes network overloading situation by distributing the load in the co-located networks. These graphs also show that with load balancing idle state of network can also be avoided in parallel with avoiding network overloading state and hence increases the available resource utilization efficiently.

5.4 Average throughput

The statistical results of the average throughput at different networks such as satellite, WiMax and Wi-Fi networks are shown in figures from Figure 14 to Figure 15. The Figure 14 is representing the graph of average throughput for each network using load balancing algorithm, whereas the average throughput of each network using non-load balancing algorithm is shown in the Figure 15. It can be easily seen from the two graphs shown in Figure 14 and Figure 15 that with load balancing algorithm all the networks showed improved average throughput as compared to the non-load balancing algorithm for RAT selection. This is because all the networks are being utilized at all the time and no network stays idle when mobile nodes are in their coverage areas.
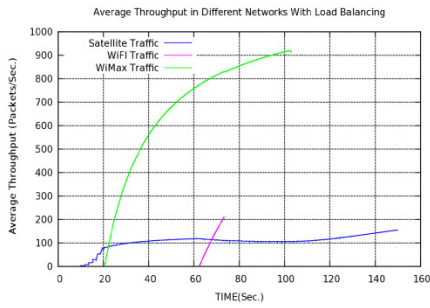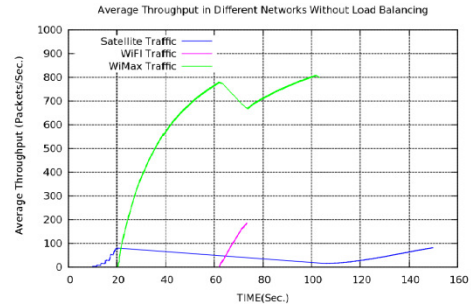
Figure 14: Avg. Throughput with LB   Figure 15: Avg. throughput without LB

5.5 Handover latencies

The handover latencies have been calculated for each scenario in the simulation by taking the difference between the time of the first packet received in the target network and the time of last packet received on previous serving network. Figure 16 shows the total handover latencies observed by each mobile user in 5 nodes scenario having different number of mobile users. The blue bars represent handover latencies observed by each node in 5 nodes scenario with load balancing algorithm andthe brown bars represent handover latencies observed by each mobile node without load balancing algorithm. It is shown in Figure 12 to Figure 13 and in Table 2 that without load balancing the networks with higher offered data rates are preferred by the mobile users and therefore networks like WiMax and WiFi become overloaded when group of mobile users enters their coverage areas. The networks with lower offered data rate are given lower priority therefore in overlapped coverage areas Satellitenetwork remains under loaded; which is poor utilization of available radio resources. As congestion at overloaded networks increases, this reduces the network availability and increases the call dropping and blocking probability.
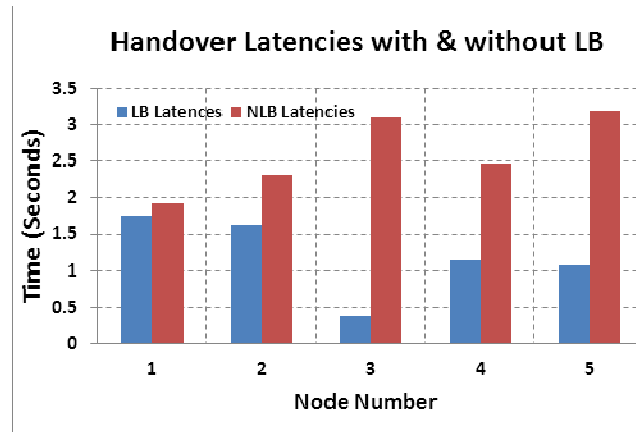

Figure 16: HO Latencies in 5 mobile users scenario

However with load balancing the load distribution remains uniform and all networks share load in overlapping coverage areas.As Figure 16 shows thetotal handover latencies observed by each mobile user in considered scenarios with load balancing is lower as compared to the scenarios without load balancing and this is because load balancing algorithm reduces the number of handovers for most of the mobile users.

6. CONCLUSION

In this paper a load aware RAT selection algorithm has been presented with the comparison of results generated by simulation scenarios using load balancing algorithm and without load balancing.Three different attributes have been compared in both types of scenarios such aswith load balancing algorithm and without load balancing.Considered attributesfor observation are load distribution on each of the available network, average throughput of each network and handover latencies. The results showed that with load balancing all three parameters showed improvement in the target heterogeneous wireless network architecture.Load balancing algorithm assures the fair load distribution between the overlapping networks whereas without load balancing different networks show abrupt load variationswhich decrease the performance with high congestion, high call dropping probability and blocking probability at overloaded network. Load balancing approach utilizes the available radio resources efficiently. Handover latencies are minimized, as it does not require all the mobile users to handover when load balancing algorithm is used. Hence the load aware RAT selection is a better approach as it offershigh radio links utilization with minimum number of handovers and hence low handover delays, minimizedcall/connection blocking and dropping probability and ability to maximize the network availability with uniformly distribution of load in co-located networks.

REFERENCES

[1]   IEEE Std. 802.21, Media Independent Handover Services, 2009.

[2]   B. Li, W. X. Shi, N. Li, "A Hierarchical Semi-centralized Architecture for Load Balancing of Heterogeneous Wireless networks", 2010 Second International Conference on Networks Security Wireless Communications and Trusted Computing NSWCTC), vol.2, pp.28–31, Apr. 2010.

[3]   Y. R. Lan, T. Yu, "A dynamic central scheduler load balancing mechanism", 1995 IEEE Fourteenth Annual International Phoenix Conference on Computers and Communications, pp.734–740, Mar. 1995.

[4]   G. Q. Ning, G. X. Zhu, "Load balancing based on traffic selection in heterogeneous overlapping cellular networks (in Chinese)", Journal of Chinese Computer Systems, vol.27, no.11, pp.2036–2041, Nov. 2006.

[5]   G. Q. Ning, G. X. Zhu, L. X. Peng, X. F. Lu, "Load balancing based on traffic selection in heterogeneous overlapping cellular networks", The First IEEE and IFIP International Conference in Central Asia on Internet, Sept. 2005.

[6]   B. Venkata Ramana, DeveshAgrawal, C. Siva Ram Murthy, "Design and performance evaluation of meghadoot – A hybrid wireless network architecture", 2006 IEEE International Conference on Networks, vol.2, pp.605–610, Sept. 2006.

[7]   P. Arun, V. Rajesh, T. Rajeev, N. Kshirasagar, "Multiple mobile routers based seamless handover scheme for next generation heterogeneous networks", First International Conference on Networks and Communications, NETCOM 2009, pp.72–77, Dec. 2009.

[8]   M. X. Li, D. L. Xie, B. Hu, Y. Shi, S. Z. Chen, "A multihop routing mechanism based on fuzzy estimation for heterogeneous wireless networks," IEEE Vehicular Technology Conference, pp.1–5, Sept. 2009.

[9]   K. Mokhtarian, M. Hefeeda, "Authentication of Scalable Video Streams with Low Communication Overhead," IEEE Transactions on Multimedia, 2011.

[10]   K. Sharif, L. J. Cao, Y. Wang, T. Dahlberg, "A Hybrid Anycast Routing Protocol for Load Balancing in Heterogeneous Access Networks," International Conference on Computer Communications and Networks, ICCCN 2008, pp.99–104, Aug. 2008.

[11] A. Zouari, L. Suciu, J. M. Bonnin, K. Guillouard, "A Proactive and Distributed QoS Negotiation Approach for Heterogeneous Environments: An Evaluation of QoS Signalling Overhead," The 2nd International Conference on Next Generation Mobile Applications, Services, and Technologies, NGMAST 2008, pp.41–46, Sept. 2008.

[12] W. Shi, B. Li, N. Li, C. Xia, "A Network architecture for load balancing of heterogeneous wireless networks", Journal of networks, VOL. 6, NO. 4, April 2011.

[13] M.Ali, P.Pillai, Y.F.Hu and K.Xu, "Load aware radio access selection in heterogeneous terrestrial wireless networks", International Journal of Computer Networks & Communications (IJCNC) Vol.3, No.6, November 2011.

[14] M.Ali, P.Pillai and Y.F.Hu, "TCP Performance evaluation over heterogeneous wireless networks using MIH", 26th International conference on CADCAM, Robotics & Factories of the Future, held on 26th -28th July in Malaysia.

[15] M.Ali, K.Xu, P.Pillai, and Y.F.Hu, "Common RRM in satellite-terrestrial based Aeronautical communication networks", Proceedings of PSATS. 2011, 328-341.

[16] M.Ali, P.Pillai and Y.F.Hu, "Load-aware radio access selection in future generation wireless networks", 4th International ICST Conference on Personal Satellite Services, PSATS 2012.

[17] E.Hossain, "Heterogeneous wireless access networks: Architectures and protocols", Springer Science & Business media, LLC, 2008, ISBN: 978-0-387-09775-3, page 204.

[18] W.Song, W.Zhuang, and Y.Cheng, "Load balancing for cellular/WLAN integrated networks", IEEE network, vol. 21, issue 1, 2007.

[19] H.Jeounglak, K.JiYeon, K.Jin-Up and K.Sang-Ha, "Dynamic load balancing architecture in heterogeneous wireless network environment", 9th International Symposium onCommunications and Information Technology, ISCIT 2009.

[20] A.Tolli and PetteriHakalin, "Adaptive load balancing between multiple cell layers", Proc on Vehicular Technology Conference, Vol.3, Fall 2002, pp. 1691 - 1695.

[21] L.Giupponi, Ramon Agusti, Jordi Perez-Romero and OriolSallent, "A Novel Approach for Joint Radio Resource Management Based on Fuzzy Neural Methodology", IEEE Tr.on Vehicular Technology, Vol. 57, No. 3, 2008, pp. 1789-1805.

[22] A.Pillekeit, FariborzDerakhshan, Bruno Muller-Clostermann, "Modeling of common radio resource management scenarios", Proc. of spring simulation multiconference, Vol. 1, 2007, pp. 205 -210.

[23] O.Holland, et al., "Development of a Radio Enabler for Reconfiguration Management within the IEEE P1900.4 Working Group", Proc. on Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySpan), April 2007. pp. 233 - 239.

[24] Shiann-TsongSheu and Chih-Chiang Wu, "Dynamic Load Balance Algorithm (DLBA) for IEEE 802.11 Wireless LAN", Tamkang Journal of Science and Engineering, vol. 2, 1999.

[25] IoannisPapanikos and Michael Logothetis, "A Study on Dynamic Load Balance for IEEE 802.11b Wireless LAN", 8th International Conference on Advances in Communications and Control, COMCON'01, June 2001.

[26] Ted Scully, Kenneth Brown, "Wireless LAN load balancing with genetic algorithms", Knowledge Based Systems, (2009).

[27] Mun-Suk Kim and SuKyoung Lee, "Load balancing and its performance evaluation for layer 3 and IEEE 802.21 frameworks in PMIPv6-based wireless networks", WIRELESS COMMUNICATIONS AND MOBILE COMPUTING, (2009, DOI: 10.1002/wcm.832.

[28] SuKyoung Lee, KotikalapudiSriram, Kyungsoo Kim, Yoon Hyuk Kim, and Nada Golmie, "Vertical Handoff Decision Algorithms for Providing Optimized Performance in Heterogeneous Wireless Networks", IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, JANUARY 2009.

[29] Alexandros Kaloxylos, IoannisModeas, FotosGeorgiadis and Nikos Passas, "Network Selection Algorithm for Heterogeneous Wireless Networks: from Design to Implementation", Network

Protocols and Algorithms, ISSN 1943-3581, 2009.

[30]   Kamil H. Suleiman, H. Anthony Chan, Mqhele E. Dlodlo, "Load Balancing in the Call Admission Control of Heterogeneous Wireless Networks", IWCMC'06, July 3–6, 2006, Vancouver, British Columbia, Canada.

[31]   Abdallah AL Sabbagh, "A Markov Chain Model for Load-Balancing Based and Service Based RAT Selection Algorithms in Heterogeneous Networks",World Academy of Science, Engineering and Technology 73 2011.

[32]   www.isi.edu/nsnam/ns/index.html.