Chapter III
QoS based handover management in 4G Networks

3.1 Introduction

2G and 3G networks, requires a type of handoff, called horizontal handoff. It requires a terminal device for changing cells among the same type of network, in CDMA network, for maintaining continuity in service. In 4G network environment, vertical handoff is provided in addition to horizontal handoff. The vertical handoff mechanism permits a terminal device to shift their networks among different types of prevailing networks (ie., among 3G and 4G networks) which is clearly transparent to end user applications. Thus, the main task is to make vertical handoffs to occur among the pairs of various types of networks like 2G, 3G, WLAN, WMAN, satellite and 4G networks. The greater challenge attributes that occur with QoS requirement is the vertical handoff. [11]

3.2 Inter Radio Access Technology [Inter-RAT]

Inter-RAT denotes Inter Radio Access Technology. RAT can represent any wireless technologies like CDMA, GSM, LTE, WiMAX or TD-SCDMA that is used for granting air interface to mobile in order to avail different services like email, voice, VOIP and internet. Inter-RAT is used for performing handoff between various RAT. MT of one RAT generates neighbouring cell measurements and forwards these reports to the network. Depending upon the measurement report as received by the mobile terminal, the network initiates handover from one RAT to other RAT. This is known as IRAT handover. Since the core network is interlocked between the RATs (WCDMA and GSM), it plays a prominent role in initiating handover among RATs. Upon completion of the handover with the new RAT, the previous RAT channels are released. For IRAT handover, inter RAT measurements are very important as even while mobile (UE) is active in one mode (or RAT), it will continuously take measurements for other RATs for handover decision. Network will also keep adding the received signal strength from the mobile for commanding UE (mobile) to perform a handover.
Generally, the inter-RAT handover consists of handover preparation phase and handover execution phase [68]. For handover that is taking place from UMTS to WiMAX, when the inter-RAT handover conditions are satisfied, RNC gives instructions to UE to enter into compressed mode of the handover preparation phase. During the compressed mode, the UE furnishes the measured data of the target network to the network with the help of Measurement Reports message. During that time, other prominent wireless link parameters like round trip time (RTT) and BDP are computed by Radio Network Controller. Following that, the inter-RAT handover will move to execution phase if the RNC performs a positive handover decision. During the handover from WiMAX to UMTS, the scanning period [69] of the network will be marked as compressed mode of WiMAX.

3.2.1 Limitations of existing Inter-RAT model

Fourth generation (4G) network comprises of various radio access technologies like GSM/GPRS, UMTS, WIFI and WiMAX. Various steps are taken to identify the existing problems pertaining to the future mobile technology. One among them is to manage and find a suitable solution for vertical handover mechanism. Various solutions related to mobility management like MIPv6/FMIPv6, SCTP and inter-RAT (Radio Access Technologies) handover are in existence[4][5]. Among these solutions, the layer 2 inter-RAT handover solution was considered as a suitable one due to very high reliability in its handover procedure. But this model supports the inter-RAT handover mechanism among cellular networks and not favouring inter-RAT handover between WiMAX and UMTS. Another limitation in providing an effective inter-RAT handover management is the interworking architecture and its coupling action.

Among various types of prevailing coupling mechanisms, integrated coupling envisage better handover performance by taking complex modification to existing network protocol mechanism. To address various above mentioned issues and to solve various existing inter-RAT handover problems such as packet loss, long handover latency and false fast retransmit, a layer 2 inter-RAT handover mechanism was developed by introducing a new common sub layer called IW (InterWorking) sublayer and SR ARQ (Selective Repeat ARQ) technique. In layer 2, a new sublayer, called IW sublayer is present on the RNC and UE and it is coupled to the top of PDCP
(UMTS) and MAC (WiMAX) sublayer. When compared to R-LLC and SDU Reconstruction which are the context transfer mechanisms, IW sublayer is predicted to achieve a perfect handover mechanism for TCP traffics with negligible loss.

Many handover issues like long handover latency, BDP (Bandwidth Delay Product) mismatch, delay spikes, packet loss, premature timeout, false fast retransmit are still prevailing. For achieving real-time and throughput-sensitive traffics, problems like packet losses and long handover latency are not preferable. In this model [70], a sub layer locating on the BTS named R-LLC (Link Layer Control) is arranged. This R-LLC sub layer plays the role of conventional LLC and once again transmits the packets that are lost during inter-RAT handover when the retransmission timer gets expired. The results of the simulation show zero packet loss for handover and cell reselection mechanisms. Still, the packet loss is indicated by retransmission timer which is set to 5 seconds for its timeout. Such a longer period is not favorable to keep TCP congestion window from getting shrunk. Also, the configuration of retransmission window is not mentioned in this model.

In order to make a handover lossless, Sachs et al [71] proposed SDU (Service Data Unit) reconstruction scheme. Here a data packet that is segmented already and stored in the PDU (Packet Data Unit) buffer of source link is first reconstructed behind SDU and then moved forward towards the target link and also to the SDUs from the SDU buffer. When these methodologies are applied during inter-RAT handover from UMTS to WiMAX or from WiMAX to UMTS for TCP traffics, the traffic performance was significant and improving. The reason being the occurrence of asynchronization of PDCP (Packet Data Convergence Protocol) PDU sequence number which took place just before a inter-RAT handover from UMTS to WiMAX. This creates asynchronization problem leading to unreliable handover procedure.

A typical problem which is caused by BDP mismatch between WiMAX and UMTS during handover from higher bandwidth data network (WiMAX) region to lower bandwidth network (UMTS) is buffered. The UMTS network undergo buffer overflow when a TCP congestion window for WiMAX is very much greater than the buffer allocation per mobile in UMTS. For SDU Reconstruction scheme, the TCP congestion window is less than the buffer size of UMTS and so the buffered packet sent from WiMAX to UMTS have the probability to
overflow the UMTS queue. This is because of more amount of packets that are buffered by the queue in WiMAX than the queue size of UMTS. This happens due to inflated transmission time.

3.3 Problem Identification

UE creates certain measurements by searching the networks range, bandwidth and quality of the service at a particular interval of time for determining link quality in a radio network. By the acquired measurement a report is created in MT which is used during handover for determining the operator and range of bandwidth to be allotted in new network. During handoff process the report created is evolved, if any new attachment is found or the old network strength is reduced. Thus, the MT handover occurs from one network to another network which has strong signal and link quality. But the MT lags in having the information required for perfect handoff causing break in network during handoff or loss of network at remote condition which cause data interruption resulting in data transfer from first packet like new transfer. [29]

Moreover existing techniques comprise of inter-RAT model on vertical handoff decision for heterogeneous networks concentrate on any one or very few of the parameters like data rate, signal strength, battery efficiency, network coverage area, network latency and velocity of MT. But for effective handoff decision and network selection, not few but all these parameters have to be considered together.

To overcome all the above mentioned issues on the handoff mechanism in 4G network, a novel method “QoS aware vertical handoff management” (QAVHD) for 4G networks is proposed. Initially, when the mobile terminal (MT) on movement finds a new network, it collects the network information of the respective network that includes network coverage area, signal strength, battery efficiency, data rate, velocity of MT and network latency. Then MT computes and compares estimated measurements with its old network and network which provides better QoS is selected as current network. The old network then transfers the information of amount of data transferred and new data to be transferred to the new network.
3.4 Estimation of Network Parameters using QAVHD model

3.4.1 Received Signal Strength (RSS)

The measurement of the power existing in a received radio signal is defined as the received signal strength. Generally, the received signal strength RSS at time $t$ is given using Eq: (3.1) [38]

$$\text{RSS} = \lambda_t S_{t-d} + \vartheta_t$$  \hspace{1cm} (3.1)

Where

- $\lambda_t$ = radio path attenuation.
- $S$ = transmitted signal waveforms.
- $\vartheta$ = noise component.
- $d$ = distance-dependant delay.

The signal strength is utilized to indicate the network availability and the existing network can be detected only if it has good signal strength.

3.4.2 Network coverage area (NCA)

The frequent handoffs cause delay and packet loss. Hence the network providing a large coverage area allows the mobile users to prevent frequent handoff while they roam. [38]

3.4.3 Estimation of Available Bandwidth (AB)

The available bandwidth (AB) at a link represents the idle capacity. This is because, at any time, the link is either idle or transmitting packets at the maximum speed at some time $T$. [72]

The available bandwidth at link $i$ at time $t$ ($\text{AB}_i(t, T)$) is given using Eq. (3.2)
\[ AB_i(t, T) = \frac{1}{T} \int_t^{t+T} (LC_i - \tau_i(t))dt \]  (3.2)

Where

- \( LC_i \) = link capacity
- \( \tau_i \) = traffic

The available bandwidth along a path is the minimum available bandwidth of all traversed links.

**3.4.4 Data rate (DR)**

For the given channel bandwidth and RSS, the maximum achievable data rate is estimated using the Shannon’s capacity theorem. [14][73] The maximum achievable data rate (DR) is given using Eq: (3.3)

\[ DR = BW \log_2 (1 + \text{RSS}/\mu) \]  (3.3)

where

- \( BW \) = channel bandwidth
- \( \mu \) = total noise or interference power over the bandwidth

The network should be selected in such a way that it transfers signals at high data rate. This minimizes the service-delivery time for non-real time service and improves the QoS for adaptive real time services.

**3.4.5 Velocity of Mobile Terminal (v)**

The velocity of the mobile terminal (v) is estimated using velocity estimation and the power spectral density of the received signal envelope (VEPSD) technique shown in Eq . 3.4 [14]

\[ v = \left( \frac{z}{c} \right) \gamma_d \]  (3.4)
where

\[ z = \text{speed of light in free space} \]
\[ \gamma_c = \text{frequency of the received signal} \]
\[ \gamma_d = \text{maximum Doppler frequency} \]

When the mobile terminal is moving at high velocity, the handoff is not encouraged as the handoff may revert back to the original network in short time. Thus the high mobile users are linked to the upper layers and get benefitted from a large coverage area.

3.4.6 Network Latency (NL)

When a mobile node moves from its existing location to the new base station range, connection with the old base station is disconnected. Then it proceeds with movement detection and address configuration that includes handoff latency which has to be minimum for seamless handoff. This is inclusive of the queuing and contention delays of the data. Thus the one-way delay should be less than 150ms which is so minimum that users cannot notice the delay [61]. The high network latency degrades the applications and the data transfer.

3.5 Estimation of MT information

3.5.1 Battery Depletion Rate (R\(_b\))

Let \( P_{tx1} \) and \( P_{rx1} \) be the transmission and reception power of network \( (N_1) \)

Let \( P_{tx2} \) and \( P_{rx2} \) be the transmission and reception power of network \( (N_2) \)

Let \( a \) and \( b \) represent the constants

When the mobile terminal uses only network \( N_1 \) (e.g. WiMAX), the power consumed is given using Eq (3.5) [74]

\[ \Delta E_{N_1} = x \cdot P_{tx1} \cdot L \cdot (1+a)/P_{rx1} \]  \hspace{1cm} (3.5)
where

\[ x = \text{number of packets} \]
\[ L = \text{packet length} \]

When the mobile terminal uses only network \( N_2 \) (e.g. WLAN), the power consumed is given using Eq (3.6)

\[
\Delta E_{N_2} = x\cdot P_{tx2} \cdot L \cdot (1+b)/P_{rx2}
\]  

When the mobile terminal performs the handover from \( N_1 \) to \( N_2 \), the battery depletion rate (\( \omega \)) is estimated using Eq (3.7).

\[
\omega = \frac{\Delta E_{N_2}}{\Delta E_{N_1}}
\]

\[
\omega = \left[ P_{tx2} \cdot (1+b)/P_{tx1} \cdot (1+a) \right] \cdot \left[ P_{rx1}/P_{rx2} \right]
\]

3.5.2 Transmission Power (\( P_{tx} \))

The transmission power of mobile terminal (\( P_{tx} \)) is the function of the received signal strength (RSS) and distance (d) among the mobile terminal and base station. [75]

\[
P_{tx} = \frac{\text{RSS} \cdot d}{\alpha}
\]  

Here \( \alpha = \text{constant} \)

3.5.3 Bandwidth Capability (\( C_{bw} \))

The measure of the width of the frequency range is termed as bandwidth. It is the difference among upper and lower frequencies in a contiguous set of frequencies. The bandwidth capability of the mobile terminal needs to be managed for offering seamless handoff for quality of service. It is computed based on the link capacity of the network. [6]
\[ C_{bw} = L_C \times \left( \frac{T_i}{T_{in}} \right) \]  \hspace{1cm} (3.9)

where

- \( L_C \) = Channel link capacity
- \( T_i \) = Idle time period in the predefined time period \( T_{in} \)

Figure 3.1 represents the 4G wireless network model. It includes the various networks like UMTS, WLAN and WiMAX. The architecture contains the respective mobile terminals (MT). The base station of the networks offers the wireless communication medium to mobile nodes.
3.6 QoS based Handoff Management

The steps involved in the QoS based handoff Management is described using the following algorithm

1) Initially, when the mobile terminal (MT) on its movement finds a new network (NN), it initiates the gathering of the network information of the new network.

2) The measurements include the quality of service (QoS) parameters such as signal strength (RSS), network coverage area (NCA), data rate (DR), available bandwidth (AB), velocity of MT (v) and network latency (NL) (Estimated in section 3.2.1-3.2.6). These parameters are measured at the particular interval of time t.

3) The measurements are consolidated into a report (MR_{new}) and maintained in MT. MT already maintains its own network information (MR_{old}). These measurements are refreshed in certain interval of time t.

4) Along with existing MR, MT also adds its information that includes the battery depletion rate (R_{b}), transmission power (P_{tx}), bandwidth capability (C_{bw}) (Estimated in section 3.3.1-3.3.3) and standby time (T_{s}) [76].

5) Following the measurements, both networks information are compared.

\[(RSS_1 \leftrightarrow RSS_2), (NCA_1 \leftrightarrow NCA_2), (DR_1 \leftrightarrow DR_2), (AB_1 \leftrightarrow AB_2), (v_1 \leftrightarrow v_2), (NL_1 \leftrightarrow NL_2)\]

Table 3.1 Format of Measurement Report

<table>
<thead>
<tr>
<th>MR_{old}</th>
<th>MR_{new}</th>
<th>MT data</th>
</tr>
</thead>
<tbody>
<tr>
<td>RSS_1</td>
<td>NCA_1</td>
<td>DR_1</td>
</tr>
</tbody>
</table>
Table 3.1 represents the measurement report maintained in MT. MR_{old} represents the network information of the mobile terminal’s current network. MR_{new} represents the measurement of new network which the MT has visited. MT data represent the information about the mobile terminal’s current network exclusive of the network information.

6) Then the network that offers better QoS is chosen as current network using the following condition.

\[ \text{If} \ (\text{RSS}_2 > \text{Th}) \ & \ (\text{NCA}_2 > \text{Th}) \ & \ (\text{DR}_2 > \text{Th}) \ & \ (\text{AB}_2 > \text{Th}) \ & \ (\nu_2 < \text{Th}) \ & \ (\text{NL}_2 < \text{Th}) \]

Then

The MT chooses NN as its current network.

End if

The accepted threshold values for the measured parameters are shown below:

i) \ RSS \ = \ High
ii) \ NCA \ = \ Large
iii) \ DR \ = \ High
iv) \ \nu \ = \ Low
v) \ NL \ = \ Low
vi) \ AB \ = \ High

7) Following the network selection, the old network (ON) transfers the information of amount of data transferred and new data to be transferred to NN without the packet loss or break in network.

ON \rightarrow NN: Data
This process is executed devoid of packet loss and network break since the network that satisfies QoS requirements are selected.

8) If ON was 4G network and MT does not find any range of 4G network on the move, the MT is forced to search for the 3G/UMTS network. Then the network adapts to 3G/UMTS network and data transfer is established.

**Advantages of QAVHD model**

- The proposed technique minimizes the packet loss.
- It improves the network throughput.
- It reduces latency.
- It prevents the network break.

**3.7 Simulation Results**

To simulate the proposed QAVHD scheme, simulation tool NS-2 [77] is employed. During simulation, the clients (SS) and the base station (BS) are deployed in a 1000 meter x 1000 meter region for 50 seconds of simulation time. It has four base stations wherein, BS1 is based on WLAN and BS2 is based on 802.16 WiMAX and remaining BS3 and BS4 are based on UMTS. Every network contains five mobile nodes as shown in figure 3.2. The transmission range of every node is 250 meters.

At 5 seconds, MN1 from BS1 begins to handoff to BS3 of UMTS network. At the same time, MN6 from BS2 of UMTS network begins to handoff to BS4 of WiMAX network. This is illustrated in Figure 3.3.
Figure 3.2: Simulation Setting of Topology for UMTS, WiMAX & WLAN (QAVHD)
Figure 3.3: Handoff Scenario for UMTS, WiMAX & WLAN (QAVHD)
The simulation settings for UMTS and WiMAX are presented in table 3.2 and table 3.3. The simulation settings used for 802.11 and general parameters are summarized in table 3.4.

Table 3.2. Simulation settings for UMTS

<table>
<thead>
<tr>
<th>Mac</th>
<th>Mac/Hsdpa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base stations</td>
<td>2</td>
</tr>
<tr>
<td>Clients</td>
<td>5</td>
</tr>
<tr>
<td>Downlink/Uplink BW</td>
<td>32kbs</td>
</tr>
<tr>
<td>HS-DSCH Scheduling mode</td>
<td>Proportional Fair Scheduling</td>
</tr>
<tr>
<td>RLC buffer size</td>
<td>10Mb</td>
</tr>
<tr>
<td>MN speed</td>
<td>10 m/s</td>
</tr>
</tbody>
</table>

Table 3.3: Simulation settings for WiMAX

<table>
<thead>
<tr>
<th>Mac</th>
<th>802.16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base stations</td>
<td>1</td>
</tr>
<tr>
<td>Clients</td>
<td>5</td>
</tr>
<tr>
<td>Physical Layer</td>
<td>OFDM</td>
</tr>
<tr>
<td>Packet Size</td>
<td>100 bytes</td>
</tr>
<tr>
<td>Frame Duration</td>
<td>0.005</td>
</tr>
<tr>
<td>Rate</td>
<td>50 to 250 kb</td>
</tr>
<tr>
<td>MN speed</td>
<td>10 m/s</td>
</tr>
</tbody>
</table>
Table 3.4: General Simulation Settings

<table>
<thead>
<tr>
<th>Area Size</th>
<th>1000mtsX 1000mts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mac</td>
<td>802.11</td>
</tr>
<tr>
<td>Base stations</td>
<td>1</td>
</tr>
<tr>
<td>Clients</td>
<td>5</td>
</tr>
<tr>
<td>Radio Range</td>
<td>250m</td>
</tr>
<tr>
<td>Simulation Time</td>
<td>50 sec</td>
</tr>
<tr>
<td>Traffic Source</td>
<td>CBR and Video</td>
</tr>
<tr>
<td>No. of CBR Flows</td>
<td>4</td>
</tr>
<tr>
<td>No. of Video Flows</td>
<td>2</td>
</tr>
<tr>
<td>Video Trace File</td>
<td>JurassikH263-256k_trace.dat</td>
</tr>
<tr>
<td>Packet Size</td>
<td>100 bytes</td>
</tr>
<tr>
<td>Rate</td>
<td>50 to 250 kb</td>
</tr>
</tbody>
</table>

3.8 Performance Metrics

The proposed QoS Aware Vertical Handoff Decision (QAVHD) is compared with the Inter-RAT model and the performance of QAVHD model is analysed in detail [29]. The performance is evaluated based on the following below mentioned metrics.

**Throughput:** It represents the amount of traffic (real time or non-real time) as received in the destination. It is represented in Megabits / second.

**Delay:** It is nothing but the average end to end delay that has occurred at the destination during the data packet flow.
3.8.1 Performance Metrics - Based on Rate

In this experiment, the rate is varied as from 50 kb to 250 kb in multiples of 50. The non real time handoff latency and NRT throughput are varied against the increase in rate. The variations are shown in graphs fig.3.4 and fig.3.5 and analysed.

Figure 3.4 shows a linear increase in Handoff Latency against the rate for both inter-RAT and QAVHD models. The handoff latency with non real time traffic has a maximum value of 2.1877 in case of inter-RAT model and has a maximum value of 1.3742 with the proposed QAVHD model. The delay is more in case of Inter-RAT model than the suggested QAVHD model. In the overall process of delay estimation, it is calculated that QAVHD model provides a less delay of 70.45% than Inter-RAT model. This shows that the proposed model QAVHD reduces the handoff latency.
Figure 3.5 shows the linear variation of rate against throughput for the Non Real Time Traffic. When the data rate is the maximum, the throughput value is around 12.505 for QAVHD model and the value is 5.7267 for Inter-RAT model. In the overall analysis, we can observe that the proposed QAVHD technique has 53% more throughput than Inter-RAT model when the data rate is increased. This is mainly because the suggested QAVHD technique provides priority in selecting the suitable bandwidth for the NRT flows.
Figure 3.6 shows a decrease in Handoff Latency against the rate in case of both inter-RAT and QAVHD models. The handoff latency with real time traffic has a minimum value of 0.2024 in case of inter-RAT model and has a minimum value of 0.2861 with the QAVHD model. The delay is more in case of Inter-RAT model than the suggested QAVHD model. In the overall process of delay estimation, it is calculated that QAVHD model provides a less delay of 19.61% than Inter-RAT model. This shows that the proposed model QAVHD reduces the handoff latency effectively.

Figure 3.6  Variation of Handoff Latency (RT) against Rate
Figure 3.7 shows the variation of throughput against rate for the Real Time traffic. The throughput increases with increase in data rate. It can be noted that the proposed QAVHD technique gives better results than Inter-RAT model by attaining 19.4% higher throughput. This is because, the suggested QAVHD technique provides transmission priority for RT flows than other models.
Figure 3.8 shows the variation of Rate against packet data delivery ratio. As the data rate increases, the delivery ratio decreases more rapidly with respect to Inter-RAT model. It can be inferred that the increased traffic rate results in decreased delivery ratio for inter-RAT model. But in case of QAVHD model, the delivery ratio is more than inter-RAT model and almost it is uniform with increase in date rate. Also, QAVHD technique yields 41% less packet drop than Inter-RAT model. Hence, it is concluded that the proposed QAVHD technique has better packet delivery ratio.
Figure 3.9 shows the variation of packet drop against data rate. It can be observed that increased traffic rate results in increased packet drop in both the cases. The increase is very high in inter-RAT model at higher traffic. But the packet drop is more in case of Inter-RAT model than QAVHD model as seen from the graph. It is seen that QAVHD technique yields 41% less packet drops than Inter-RAT. Hence, it is concluded that the proposed QAVHD model is a better model for controlling packet drop.
3.8.2 Performance Metrics - Based on Time

This part of the work explains the analysis of metrics for different time variants.

Figure 3.10 shows the variation of latency against handover time for the NRT traffic. As time increases, the handoff latency increases steeply with inter-RAT model whereas it shows less value with QAVHD model. It is observed that QAVHD technique is better than Inter-RAT when the time is increased. This is because our QAVHD technique provides bandwidth priority for NRT flows.
Figure 3.11 shows the variation of time against Throughput (NRT). The throughput increases linearly with increase in time in both the cases. Throughput value is much better in QAVHD model than inter-RAT model which suggests that QAVHD model is better as seen in the above graph.

Figure 3.11 Variation of Throughput (NRT) against Time
Figure 3.12 depicts the variation of handoff latency with respect to time (Real time). The handoff latency varies linearly with increase in time in both the cases. The handoff latency is less in QAVHD model compared to inter-RAT model. Hence it can be concluded that when compared to inter-RAT model, the performance of QAVHD model is better in terms of handoff latency as seen from the above graph.
Figure 3.13 shows the variation of throughput against time for the Real Time traffic. As the time increases the throughput value increases in both models up to certain time value and beyond certain limit, the variation is maintained and almost uniform with respect to increase in time.

It can be observed that our QAVHD technique yields better results than Inter-RAT when the time is increased. The throughput value is higher in QAVHD model than the results obtained from inter-RAT model. This is because our QAVHD technique provides transmission priority for RT flows.
Thus, in this work, we have proposed a QoS based handover management in 4G networks. Initially, when the mobile terminal (MT) on movement finds a new network, it collects the QoS information of the respective network that includes signal strength, network coverage area, data rate, available bandwidth, velocity of MT and network latency. Then MT compares estimated measurements with its old network and network which provides better QoS is selected as current network. The old network then performs the data transmission to the new network. By simulation results, we have shown that the proposed QAVHD approach is a better model and enhances the network throughput and minimizes the latency.