PRioritized Multimedia Adaptation Scheme over Two-Hop Heterogeneous Wireless Networks (PRiMA)

Ankit Maheshwari, Shruti Agrawal, Hrishikesh Venkataraman, Member, IAENG, and Gabriel-Miro Muntean

Abstract – Live multimedia streaming is one of the greatest promises of a wireless network operator. In recent years, there has been an upsurge of interest in the feedback-oriented multimedia streaming in the wireless domain, in both industry and academia. However, the lack of an acceptable and guaranteed quality of service (QoS) in the wireless domain results in the media packets experiencing dynamic variations in bandwidth, delays and loss rate as they traverse from the sender to the receiver. Moreover, the plethora of wireless devices in the market ranging from i-phone and PDA to laptop, HDTV etc. makes it difficult to provide optimum video quality of the same multimedia content to all the devices simultaneously.

In this work, a WiMAX-based two-hop cellular network is considered for multimedia transmission across different hand-held wireless devices. A feedback-based prioritized multimedia adaptive scheme (PRiMA) is proposed for two-hop heterogeneous wireless networks, which ensures that different clients receive the same video with different perceived quality (satisfactory, good and excellent); depending on: (a) the action content in the video, (b) the end-user’s choice of service, and importantly, (c) the user’s device characteristics. The base station (as the web server) communicates with the end-users through the relay nodes which act as proxy servers. It has been observed that such a prioritization not only results in a higher average perceived quality of the network but also provides a higher perceived quality to majority of the end-users, as compared to the case when there is no prioritization.

Index Terms -- Prioritization, proxy-client-server (PCS), quality-oriented adaptive scheme, two-hop.

Manuscript received Feb 8, 2010.

A. Maheshwari is with Next Generation Wireless Systems, Dhirubhai Ambani Institute of Information and Communication Technology (DAIICT), Gandhinagar, India. Email: ankit_maheshwari@daiict.ac.in.

S. Agrawal is with Next Generation Wireless Systems, DAIICT, Gandhinagar, India. She did this work while working in the Performance Engineering Laboratory (PEL) at Dublin City University (DCU), Ireland, as part of her research work. Email: shruti_agrawal@daiict.ac.in.

H. Venkataraman is with Next Generation Wireless Systems, DAIICT, India. He is presently a research fellow at PEL, DCU, Ireland. Email: hrishikesh@eeng.dcu.ie; Phone No: +353-1-7-6401 Fax No: 353-1-7005508

G. M. Muntean is with School of Electronic Engineering and Co-Director of PEL, DCU Ireland. Email: munteang@eeng.dcu.ie; Phone No: +353-1-7-6401

I. INTRODUCTION

Over the last decade, there has been a tremendous growth in the telecommunication world; particularly efforts have been made in providing voice and data services anytime, anywhere. With the arrival of internet over hand-held devices, there have been significant efforts in the recent years [1] to provide on-demand-based access to rich media and very high quality multimedia to home residences via an all-IP infrastructure [2], [3]. The current demand is for live multimedia streaming and video broadcasting over a hand-held wireless device. However, triple play services (voice, data and video) are yet to be fully implemented and deployed in the market.

There are significant technological bottlenecks that hinder its deployment in the wireless world. There are many challenges existing in the design of video streaming systems. Firstly, the time delay is usually very high, thereby causing great difficulty in real-time video broadcasting. Secondly, the power required at the hand-held device for multimedia transmission is very high. In case of video broadcasting, the battery does not last for more than a few hours. Thirdly, the wireless channel changes rapidly, when the distance between the source and the destination node is high. This would in turn require huge computations in order to provide efficient multimedia delivery, which again causes reduction in the battery power. Most important is the End-User Perceived Quality (EUPQ) which must be above a satisfactory level and must not degrade during the transmission period. Also, the varying medium makes the estimation and calculation of packet loss difficult. Hence, alternate techniques have to be adapted in order to have high quality multimedia transmissions in the wireless network.

The integration of multihop design into the conventional hierarchical wireless networks is one of the most promising architectural upgrade to meet the
next generation demand of multimedia transmission in cellular networks. In such a design, the base station (BS) (associated with a web server) communicates with the end-users in multiple hops, through intermediate relays. The BS communicates with the far-off wireless terminals through these relay nodes. This results in a shorter transmission distance, and thereby less transmit power for the transmitter, which in-turn results in less interference and importantly, higher data rates. Also, with the existence of different kinds of wireless networks in the current scenario (GPRS, CDMA2000, UMTS, WiMAX, WLAN, etc.); there is a necessity to inter-operate over different networks. For example, the end-user connected over UMTS could seamlessly switch over to WLAN or even WiMAX when it enters a campus/semi-indoor environment. The communication over heterogeneous wireless networks can be better understood by categorizing them as communication over multihop wireless networks; wherein the communication across different hops takes place over different networks. Hence the multihop design can efficiently model the multiplicity in wireless networks. In addition, the multihop design also increases the coverage area of the wireless network [3].

In the existing multimedia transmission systems, every end-user is assumed to have same characteristics which results in equal distribution of bandwidth and hence same bitrate to all the clients resulting in unbiased end-user perceived qualities. The client characteristics are not only related to the type of video that has been requested by end-user, but also depends on two other major factors: 1) the type of service the end-user chooses for multimedia transmission and 2) the resolution of the end-user device. Hence, there is a need to introduce prioritization in the wireless networks in order to take the end-users’ characteristics into account before allocating the network resources to them. This would result in a dynamic and importantly, a customized distribution of the network bandwidth, which will ensure that each client gets a bitrate which is proportional to its characteristics.

In the prioritized network, a client demanding higher quality of service (QoS) has to be served with higher priority than the one demanding for lower quality of service, which will be reflected by the amount of money the end-user pays. For instance, a high priority client (client who demands higher QoS) watching a high motion video on a 15” laptop will require far more bitrate than a lower priority client (client who demands lower QoS) watching a low motion video on a mobile phone. Similarly, a client viewing a medium action video on a smart phone will require more bitrate than a client watching the same video on a mobile phone, if both choose the same type of service. In congruence with the above mentioned logic, a medium priority client watching a low-action video on an HDTV would require more bitrate than a high priority client watching a medium-action video on a PDA.

The paper is organized as follows: Section II describes the motivation and the related work in this area. Section III describes the proposed prioritization mechanism – PRiMA and the proposed service-based, device-based and video-content based classifications. Section IV explains the function and architecture of PCS. Bandwidth allocation mechanism and the feedback scheme used are illustrated in the Section V. The simulation model is described in Section VI. The results are provided in Section VII, while the conclusions and the possible future work are addressed in Section VIII.

II. RELATED WORK

There has been considerable research done in the area of multihop networks. In [4] it is shown that for up-to five hops, the spectral efficiency is increased as compared to single hop transmission. Notably, in their landmark paper, Gupta and Kumar [5] demonstrated that the multihop design with the presence of relays significantly increases the capacity of the wireless network. Several architectures and mechanisms have been proposed in the recent past [6], [7] in order to efficiently design a Multihop Cellular Network (MCN). The authors in [8] proposed three simple routing techniques for increasing the data rate in a multihop cellular network. In addition, a time division duplexing (TDD)-based MCN offers the potential to integrate various two-hop infrastructure-based wireless networks, such as UMTS or CDMA 2000 (3G networks), WiMAX and WiFi [9]. However, resource allocation is a very challenging issue and is proved to be an NP-hard problem [10]. Hence, the focus of researchers across the world mainly has been on two-hop cellular networks. Recently, a cluster-based architecture has been proposed in [11] for two-hop cellular networks. The architecture provides an increase in the data rate without an increase in the power requirement. The frequency reuse in the cellular network is also increased. Hence it enables high quality multimedia transmission, without losing

(Advance online publication: 13 May 2010)
the battery power significantly. Several algorithms have been recently proposed for two-hop cellular networks in [12], [13] and [14], wherein, the server communicates with the nearby clients in one-hop; whereas it communicates with the far-off clients in two-hops. It has been shown that in a two-hop network the EUPQ of the video is more when compared with that obtained from an equivalent solution in a single-hop network. Extensive research has proposed various solutions for providing certain levels of quality of Service (QoS) while streaming multimedia over IP-based networks [15], [16], [17]. The main focus of these works has been on a congestion control approach to QoS, rather than increasing the EUPQ. EUPQ can vary severely in wireless environments due to fluctuations in network conditions, especially when multiple clients require streamed multimedia simultaneously.

In a wireless environment, the unpredictability and the constantly varying nature of the wireless channel necessitate the implementation of the feedback-based quality oriented adaptive scheme (QOAS) [16], [18]. The adaptive multimedia streaming solution maximizes the end-user perceived quality in highly variable and increasingly loaded network delivery conditions [17]. However, QOAS does not assign any priority to any specific users in the heterogeneous network. There is not much work done on prioritizing the clients during multimedia streaming in most of the works ([16], [17], [18]) except the work on priority-based adaptive scheme for multimedia delivery over wireless networks in [15]. However, in [15] the problem is approached in a very simplistic manner, as it does not consider any classification based on the video content nor does it addresses any form of heterogeneity among the end-users.

The given diverse nature of today’s wireless devices, end-users need to have different priorities based on their device characteristics, i.e. the screen size and resolution, whether they support full color or reduced grayscale, the amount of available local memory and the CPU power (hardware variations), application level data encodings that the client can handle given the processing and display capabilities of the end-service (software solutions). In addition, the priorities need to be assigned based on the client’s requirement of the perceived quality and willingness of end-user to pay for that particular video content. Also, given the constantly varying nature of the wireless network there is a need of an efficient feedback mechanism for ensuring adaptive streaming. Hence, in this paper, a novel prioritization scheme is proposed for efficient multimedia transmission. This is combined with QOAS to form the feedback enabled, quality-controlled prioritized adaptive multimedia – PRiMA scheme.

III. PRiMA

The goal of developing PRiMA is to have a resource-based adaptive scheme for multimedia streaming which would provide a fair quality of experience (QoE) to different users based on their characteristics and other subjective priorities.

![Fig. 1: System Design for client-server based PRiMA transmission technique](image-url)
This work proposes a prioritized adaptation scheme which allocates the network resource adaptively during streaming on the basis of the priorities (the device-type, the video content requested and the service requested), statically assigned by proxy-client-server as per the client’s needs. The scheme adjusts the stream’s bitrate to suit available network conditions which in turn affects EUPQ when streaming multimedia. The scheme is based on a client-server based feedback mechanism which enables the adaptation to be performed while the transmission is going on and also before the start of the transmission. PRiMA follows the basic QPAMS architecture proposed by the same authors in [19].

The system architecture of PRiMA is illustrated in Fig. 1. A key component at the client side is the multimedia delivery monitoring module that monitors the delivery in terms of multimedia quality according to client priority. The key component at the server-side is the PRiMA module which performs the adaptation and prioritization. The PRiMA at the server then decides the extent of adaptation, and also the category of prioritization for each user, depending on his/her demand, and the availability of the bandwidth resource.

A. Adaptive Multimedia Streaming

In order to support live multimedia streaming and video broadcasting in wireless network with increased user QoE, an adaptive client-server-based feedback approach is required, wherein, the client monitors the transmission and user QoE-related parameters, and sends them as feedback to the server which in turn adjusts the video transmission rate. In this context the state-of-the-art solution is QOAS. With QOAS client monitors the transmission and user QoE-related parameters, and sends feedback to the server which according to an adaptive mechanism adjusts the video transmission rate to the reported delivery conditions. QOAS is based on the fact that random losses have a greater impact on the perceived quality than a controlled reduction in quality [20].

QOAS adjusts the content as well as the transmission rate, increasing or decreasing the quantity of streamed video data by dynamically adjusting its quality [18]. This is done according to feedback information received from the client who bi-directionally exchanges video data and control packets with the server. During transmission the server dynamically varies its state according to the client-reported stream quality. For example, when the client reports a decrease in end-user quality, the server switches to a lower quality state, which reduces the quantity of data sent. In improved viewing conditions, the server gradually increases the quality of the delivered stream using the Quality of Delivery Grading Scheme (QoDGS) [17]. QoDGS regularly computes the quality of delivery scores, which are sent as feedback to the server. The QoDGS takes into account the end-user quality as measured by the moving pictures quality metric $Q$, which maps the joint impact of bitrate and data loss on encoded video streams quality onto the ITU-T R P.910 five-point grading scale [21].

B. Prioritized Multimedia Transmission

The multimedia transmission can be prioritized in many ways based on the factors that affect EUPQ. In this work, the users are categorized depending on the service requested by them for viewing the video stream, the resolution of the user device and the action content of the video stream. The concept of the three types of prioritization techniques is explained further.

i. Service based classification

For a given video stream, a user demanding higher perceived quality requires higher amount of resources as compared to a user seeking lower perceived quality, according to (1). Hence, the user demanding higher quality has to pay more and vice versa. The clients have been prioritized in $M_C$ categories based on the service chosen, which is categorized based on perceived quality, classified as in the ITU-T R P.910 five-point grading scale [21]. The clients are prioritized by the service chosen, incorporated by allocating biased resources. Such a prioritized video transmission technique enables the network operator to bring in more revenue by providing an adaptive service to the end-users which is directly proportional to the amount the client is willing to pay for a particular video transmission.

ii. Device based classification

Heterogeneity in the end-user devices in the network can be further used for prioritization. Prioritizing the devices is an important aspect, as in modern circumstances there exists a lot of diversity and variety in the end-user devices used. Not all devices are of the same size, with same features or same resolutions. Hence transmitting the streamed video with the same rate on all the devices will result...
in an unacceptable $Q$-value for devices like laptops etc. and a wastage of the additional bit-rate for devices like PDA’s, i-phones etc. This unbiased distribution leads to the partial wastage of network resources. Hence to incorporate this diversity and to introduce bias among the end-users in terms of the device used, the end-user devices have been classified into $M_D$ categories based on resolution of the devices. For example, the devices can be categorized in certain ranges based on resolution like $X - 2X$, $2X - 4X$ etc., where $X$ would be the minimum device resolution.

iii. Video-content based classification

Categorization of the streamed video is done keeping in mind the fact that not all videos have the same degree of animation and visual effects. In practice, there are different kinds of multimedia programs which differ in their temporal complexity. Temporal complexity implies the number of changes in the pixels between consecutive video frames. Every multimedia stream has different characteristics in terms of action sequence and rapid movement of pictures. Different video classes like news, a daily soap opera and a sports event can be incorporated in different video classes like low, medium and high action videos respectively. This categorization will be considered when the user chooses the video to be streamed and accordingly the constants in the $Q$-value calculations are chosen. Video classification is done using mpeg_stat, an MPEG analysis tool [22], and the focus is to keep ‘news’, ‘daily soaps’ and ‘blockbuster movie’ or ‘sports’ in different categories using Interpolated Macro-blocks concept [23], [24]. The tool analyses a .mpg stream and outputs the number of coded and skipped macro-blocks in a stream. The coded macro-blocks imply the number of blocks in a frame that are not matching with the previous frame and need to be coded, and those that match with the previous frame are skipped blocks. Hence by comparing the number (percentage) of blocks that have to be coded, one can measure the amount of motion content in the video streams which can be used to classify the video based on animation content. Hence, the video streams are classified into $M_V$ categories based on their motion content.

Therefore, PRiMA enables prioritized video transmission to the clients based on the action content of the video stream, service opted and screen-resolution of the device held by the user. It attempts to allocate optimum bandwidth to simultaneous users by seamlessly altering the bitrate according to the received feedback. PRiMA would not only customize the resource allocation according to client priorities, but also ensure maximum utilization of available resources.

IV. PROXY-CLIENT-SERVER: PCS

There might be multiple clients requesting for the same video which will require multiple encoding of a single video stream, thereby creating unnecessary load on the base-station. For example in a single-hop prioritized network, implementing prioritized scheme would imply, the video stream encoded once on the request of a client can be transmitted to another location (in this case PCS) where it can be stored [25]. When another client requests for the same stream, instead of asking the server to encode the video stream again, the hop can transmit the same. Therefore, a two-hop wireless network has been established as shown in Fig. 2, where the hop can be considered as a Proxy-Client-Server (PCS).

![Fig. 2: A two-hop cellular network with the relay node acting as both database server and proxy client](image-url)
A PCS with enough buffer-space is considered in the system design with enough memory space to store the data sent by the server (base station). It also acts as a database server. Each PCS serves multiple clients. QOAS-based multimedia transmissions are set-up between the PCS and all the clients. Hence, the adaptation strategy is different for each end-user, and is decided by the PCS. The PCS therefore performs seamless resource allocation to the clients on the basis of their priorities and feedback received from them.

A. Architecture of PCS

In the proposed scenario, the client will send its priorities (the service requested and device-type) and the type of video it needs to the PCS. The PCS will request the video from the main server and will also calculate the share of bandwidth (bitrate) that will be allocated to the respective client based on the priorities. This allocation will directly be related to the service opted by the client and the quality it receives. The framework incorporated in the PCS is described in Fig. 3. The video stream is assumed to be available at the PCS, which is divided into several chunks of multimedia. The size and length of each chunk depends upon the total size, length and the type of video to be streamed. ‘A’ number of chunks are formed from each video, such that duration of each chunk is fixed to a pre-defined value. Each of these video chunks is passed into the system at regular intervals. Firstly, each chunk of video stream is classified into high, medium or low action video content by the video classification methodology as illustrated in Section III. B-iii. This classification is done in the ‘classification based on video content’ block which gives the video priority (V) as output. The type of service selected by the client for watching the video has been called its client priority (C) and the priority of the device used by the end-user (based on its screen-resolution) has been named device priority (D). These priorities (C and D) are chosen by the client when it demands for the video. The information of the client i.e. the values of C and D are received by the PCS in the feedback packets sent by the client during establishment of wireless connection with it.

In the next block, i.e. the ‘set encoding rate’, a range of bitrate is set according to the video content, V. This range consists of all the possible bitrates a client can receive for viewing the particular stream according to the cost-priority chosen and the device-type of the end-user.

According to the values of C, D and V, the parameters required for the video-transmission, i.e. frame rate, bit depth and bitrate; are set in the ‘set parameters for client prioritization’ block. This block takes the priorities of the client as well as the video type as input. Based on these values, it computes the most appropriate bitrate that can be allocated to the particular client, estimating the availability and attributes of other simultaneous clients as well. This most appropriate bitrate value is a function of C and D given V, and is thus represented as f(C/V, D/V). It is sent as input to a function ‘F’, which selects the maximum possible bitrate value, from the range of values as given by the ‘set
encoding rate' block, which is closest to the most appropriate bitrate. The therefore selected bitrate is the optimum bitrate at which the client can be served, given the network resources and number of clients present. The function \( F = f(V, D, C) \) is a multivariate function of \( V, D \) and \( C \), collectively called as the client characteristics.

V. BANDWIDTH ALLOCATION AND FEEDBACK MECHANISM

A. Bandwidth Allocation

In a non-prioritized transmission technique, all the clients requesting the server simultaneously are treated equally. If \( B \) is the total bandwidth and \( N \) is the number of simultaneously communicating clients under the PS, then the bandwidth allotted to each client in case of equal treatment would be \( B/N \). However, in case of a prioritized transmission scheme, clients with different priorities would receive different bandwidths. If \( r \) is the priority factor for a client, then the bandwidth for the communicating client is \( w = B/N \times r \). In case a client has higher priority, then the average bandwidth resource factor given to a user is, \( r > 1 \), and in case, a client has lower priority than the average bandwidth resource factor given to a user is, \( r < 1 \). It should be noted that the maximum average value of the factor, \( r \) over \( N \) users is one, implying that the maximum available bandwidth, \( B \), is utilized by the system. The end-user quality is computed using the multimedia perceived quality metric proposed in [26] and expressed using the ITU-T R P.910 five-point scale for grading subjective perceptual quality [21]. If the total users are classified into \( M \) categories according to the prioritization schedule, then the perceived quality of the \( i^{th} \) category would be given by:

\[
Q = Q_0 + \chi_Q \times \left( \frac{R_i}{\gamma R} \right)^{-\frac{1}{\gamma R}} + \chi_L \times R_i \times PLR
\]  

(1)

It can be seen that \( Q \) of a particular category depends on both the packet loss rate (PLR) of the channel and the mean bit rate, \( R \). The bit rate of a user in the category \( i \) would be given by:

\[
R_i = w_i \times n
\]

(2)

\[
= \frac{B}{N} \times r_i \times n
\]

(3)

where \( n \) is the number of bits/symbol which depends on the modulation technique used, and \( r_i \) is the respective priority factor. In the prioritized technique, the bandwidth ratio allotted to each category of users not only determines the perceived quality of users in each category, but also plays an important role in determining the average perceived quality of the network. If \( Q_1, Q_2, ..., Q_M \) are the average perceived quality of the users in each of the \( M \) categories, and \( U_1, U_2, ... U_M \) are the number of users in each category, then the average perceived quality of the network is given by:

\[
Q_{avg} = \sum_{i=1}^{M} U_i \times Q_i = \sum_{i=1}^{M} \sum_{j=1}^{U_i} Q_{i,j}
\]

(4)

where \( Q_{i,j} \) is the perceived quality of the \( j^{th} \) user in the \( i^{th} \) category. It should be noted that

\[
\sum_{i=1}^{M} \sum_{j=1}^{U_i} = N
\]

(5)

where \( N \) is the total number of clients served by the proxy server. All the clients belonging to the same category is allotted the same bandwidth ratio. The PLR, the number of bits/symbol and the bandwidth of the system are usually fixed for a system. Hence, it can be seen from (1), (3) and (4) that for a given number of communicating clients \( N \) served by a single PS, the average perceived quality of the network, \( Q_{avg} \), is a complex non-linear function of the number of users in each category, \( U_i \), and the bandwidth ratio assigned to each category, \( r_i \), i.e.,

\[
Q_{avg} = f(U_1, r_1, U_2, r_2, ... U_M, r_M)
\]

(6)

In order to assess the performance of the prioritized technique in PRiMA, the number of users served by the proxy server simultaneously, i.e. \( N \), is kept constant and the number of users in each category is varied dynamically over different possible combinations.

The average perceived quality of each category and of the entire network is determined through the algorithm as described in the next section. The algorithm calculates optimum bitrate assigned to all clients based on their characteristics, i.e. values of \( C \), \( D \) and \( V \); in order to serve them with the quality of service in a particular desired range. Once the bitrates
are determined for all the clients, the EUPQ and average perceived quality can be calculated using (1) and (4) respectively.

### B. Algorithm

The clients are streamed initially at the optimum bitrate which is calculated in two steps as illustrated below.

**Step 1: Calculating bitrate ratio according to client priorities**

In the first step, the bitrate ratio for $N$ simultaneously served clients is calculated considering the client characteristic $C$. As mentioned in Section III. C-i, clients are divided into $M_C$ categories on the basis of $C$-value. The ratio in which the total bitrate is distributed amongst the clients is given as $b_{C_i}$ which satisfies

$$
\sum_{i=1}^{M_C} M_{C_i} \times b_{C_i} = B
$$

**Step 2: Re-calculating bitrate ratio by incorporating device priorities**

In the second step, the bitrates are re-calculated considering the device type possessed by the clients. As mentioned in Section III. C-ii, clients are divided into $M_D$ categories on the basis of $D$-value. The bitrate is re-distributed amongst the clients with a particular $C$-value depending on their $D$-values; the bitrate ratio therefore becomes $b_{D_i}$ which satisfies

$$
\sum_{i=1}^{M_D} M_{D_i} \times b_{D_i} = b_{C_i} \times B
$$

$$
\frac{1}{B} \left[ \sum_{j=1}^{M_C} M_{C_j} \times \left( \sum_{i=1}^{M_D} M_{D_i} \times b_{D_i} \right) \right] = B
$$

The optimum ratio of bitrate at which the clients are streamed initially, i.e. $b_i$, is given by the product of the two ratios, i.e. $b_{C_i} \times b_{D_i}$. Hence, the allocation of optimal bitrate satisfies (9).

### C. Role of Feedback

Feedback from the client is an important aspect of the multimedia streaming system. It is possible that because of congestion or increase in number of clients, the traffic in network might increase and hence with the bandwidth remaining constant, the quality of perceived video may decrease. Hence feedback mechanism enables the client to inform the server about the network state and the proxy-client-server can take necessary adaptive steps. In the proposed feedback mechanism, a client sends feedback to the PCS every time it receives chunks of data. The PCS then increases or decreases the bitrate of transmission by a small value based on the feedback. Since, the focus of this work is on end-user perceived quality; the server quantitatively decides the measures that are to be taken to ensure that the $Q$ values are in accordance with the priorities assigned in the network, based on the feedback information received by all the clients (being served simultaneously).

Feedback from client has many benefits. The PCS has the information of all the clients served at a time, which enables it to efficiently allocate the resources amongst them complying with the network constraints. If a client receives poor quality initially when the streaming starts, the quality would keep on improving based on negative feedbacks and after a few iterations it would start receiving better quality. This feedback mechanism enables a seamless adaptation of quality, by increasing/decreasing the bitrate by a small fraction of the previous available bitrate.

### D. Proposed Feedback Scheme

It is assumed that the client sends the information on bitrate received by him, to the PCS (after every fixed interval of time) when end-user starts receiving packets. According to the feedback the bitrate at which the video is being streamed to the particular client is modified, relatively with the other existing clients. By this feedback mechanism, the prioritized clients can receive a better quality, and the bandwidth would be efficiently used. The clients receive the video at an initial bitrate $b_i$, satisfying (10) with maximum prior client receiving maximum bitrate.

$$
\sum_{i=0}^{N} b_i \leq B
$$

In case of clients having the same $C$ but a different $D$, the bitrates are further manipulated, allocating maximum bitrate to the device with maximum resolution and vice versa. Then, as per the
feedback sent by the client, the corresponding bitrates are varied. If the bitrate received by lowest priority client is greater than a pre-determined threshold $Q$-value (required to view the content quite clearly) then it is decreased by a small variable factor, $\varepsilon$. There is a seamless transition in the bitrate allotted $b_i$, which varies from initial value to a most suitable value $r_i$, where, $b_i - \varepsilon : r_i : b_i + \varepsilon$ and $0 \leq \varepsilon \leq 0.5$.

At the expense of lower priority client, the higher priority clients receive a higher bitrate as compared to that they were receiving in the previous iteration. The increase in bitrate depends upon the previous bitrates of respective clients. There is a seamless transition in the bitrates received by the clients during the establishment of wireless connection and transmission, which continues till either the client with highest priority, obtains a quality corresponding to its upper threshold $Q$-value, or the clients with lower priorities obtain the lowest possible $Q$-value for observing the video. This work proposes the prioritization scheme based on theoretical simulations, which are intended to be tested in future using a WiMAX access point, as for instance, Motorola WiMAX Access Point (WAP) 400 [27]. The feedback mechanism aims at serving all the clients as per their priorities, hence further introducing biasing between them based on the respective priorities.

VI. MODELLING and SIMULATION

A. Simulation Scenario

For the simulations, $N$ is considered to be 6, i.e. 6 clients are communicating with the PCS simultaneously. In order to assess the performance of the prioritization scheme, the number of users served by the proxy server is kept constant at 6. Also, the clients, device-types and video-streams have been categorized in 3 levels, i.e. $M_C = 3$, $M_D = 3$ and $M_V = 3$ has been taken. Though the classification has been done in 3 categories, but it can be further fine grained as need arises. Clients are classified based on the service selected, and the levels are named as:

- **C1, premium service**: The end-users opting for premium service require very high bitrate as EUPQ received by them is desired to be greater than 4. Users falling in this priority level should accordingly pay the highest amount in order to receive the excellent service.

- **C2, economy service**: The end-users opting for economy service require high bitrate as they receive EUPQ received by them is in range 3-4. These users should pay a nominal amount for receiving good service.

- **C3, free service**: The end-users opting for free service require low bitrate as EUPQ received by them is in range 2-3. These users receive satisfactory service. Hence they may or may not be paying for the service, depending on the operator.

Similarly the devices are prioritized based on their resolution. This prioritization scheme, categorizes the devices into three broad categories, laptops (high resolution), PDA’s (medium resolution) and mobile phones (low resolution). The priority levels are named as:

- **D1, 2X - 4X**: This category comprises of the end-user devices with the highest resolution. This class usually includes devices like laptops etc. and require higher bit-rates.

- **D2, X - 2X**: This category comprises of end-user devices with moderate resolution. This class usually includes devices like PDA’s etc and require average bit-rates.

- **D3, till X**: This category comprises of the end-user devices with least possible resolution. These devices require comparatively much lower bitrate and include devices like common 3G mobile phones, where $X = 352 * 240$.

In terms of temporal complexity (as described in Section III. B-iii), the video programs are usually classified into 3 sections:

- **V1, high-action**: This category comprises of action movies or live sports events. These video sequences incorporate lot of movements, and hence, require a very high bit rate. They are classified as video sequences with more than 70% interpolated macro-blocks.

- **V2, medium-action**: This class includes general video programs/ drama scenes where the number of scene changes per second is higher than the low-action category. These are classified as video sequences having interpolated macro-blocks between 35-70%.

- **V3, low-action**: This includes sequences with little or no movement in the background. There is very little difference between the subsequent frames and the bit rate requirement is low. They are the video sequences with less than 35% interpolated macro-blocks.
B. Simulation Setup

The hierarchical multimedia network is established with a single server and a single PCS which serves 6 clients. The video streams are categorized into three different sections, based on their temporal complexity, i.e., $M_r = 3$. The topology proposed assumes a bandwidth of 5 MHz and a constant delay of 2 μsec. In each of the envisaged scenarios 95-99% of the total available bandwidth is used for video and multimedia communication and the remaining 1-5% is reserved for feedback purpose. A constant PLR of $10^{-7}$ is assumed throughout the analysis. Similarly, a constant transmission delay of 10 ns is assumed between the PCS and the end-users.

An important factor that bitrate $R$ depends on is the number of bits/symbol, $n$. In case of QPSK modulation technique, $n = 2$, whereas in case of 8-PSK modulation technique, $n = 3$. A higher modulation technique requires higher SINR (signal to interference noise ratio) at the receiver of the communicating link, but at the same time would result in higher $Q$. In this simulator, the video streams are assumed to be modulated with QPSK modulation technique, i.e., $n = 2$. The PCS starts streaming multimedia content to the clients according to their $C$ and $D$ values with optimum bitrate being in ratio 1.5:0.9:0.6 (as explained in Section V. A). The video stream is analysed in the PCS in chunks of 20 sec duration.

The range of bitrate that can be allocated to a particular client, as computed by the set encoding rate block (described in Section IV. A), is depicted in Table I. These bitrate values include the bitrates between the maximum and minimum bitrate required by a client with a particular combination of $C$ and $D$, which ensures that the clients receive the maximum EUPQ, given the network conditions. The range of values as shown in Table I is computed by considering the bitrate values given by the graph shown in Fig. 6. The quality received corresponding to the given bitrates is in accordance with the ideal $Q$ ranges as given in Table 2.

It has been noticed that $Q$-value changes rapidly with even a slight change in bitrate for $Q < 2$; moderately for $2 < Q < 3$; and gradually for $Q > 3$. Hence the initial ratio of bitrate for client with priority $C_1$ is kept very high as compared to that of $C_3$. The clients send feedback to PCS as soon as they start receiving chunks of video-stream. The PCS calculates $Q$-value for every client after each transmission and compares this value with the ideal range of $Q$-values with respect to $C$ and $D$. The threshold considered for ideal $Q$ for the clients in different categories according to the video content is illustrated in the Table II. If the bitrate of any client is not in the ideal range, the PCS then adapts the bitrate accordingly, by decreasing the bitrates of clients with priority $C_2$ and $C_3$, and increasing that of $C_1$.

The variation in bitrate changes $Q$-value by the relation given in (1). Therefore, the PCS allocates optimum bitrate for required service to all the clients by seamless adaptation of bitrate. Since, the $Q$-values of clients with priorities $C_2$ and $C_3$ were initially out of the required range in Table II, the bitrates for these clients is decreased by 1%, and that of clients with priority $C_1$ is increased by 2%, after every iteration.

The above process continues till clients with priorities $C_2$ and $C_3$ have the perceived quality in the ideal range. The bitrate has been modified as per the feedback received by client in a seamless manner, according to its priority. Fig. 4 compares the bitrate

<table>
<thead>
<tr>
<th>Video type</th>
<th>Ideal $Q$-range for clients with different priority</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$C_1$</td>
</tr>
<tr>
<td>High-action</td>
<td>4.2</td>
</tr>
<tr>
<td>Med-action</td>
<td>4.0</td>
</tr>
<tr>
<td>Low-action</td>
<td>3.7</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Device Priority</th>
<th>Client Priority</th>
</tr>
</thead>
<tbody>
<tr>
<td>High-action video</td>
<td></td>
</tr>
<tr>
<td>$D_1$ (Above 4X)</td>
<td>5000-7900</td>
</tr>
<tr>
<td></td>
<td>3150-7100</td>
</tr>
<tr>
<td></td>
<td>1610-5500</td>
</tr>
<tr>
<td>$D_2$ (2X - 4X)</td>
<td>4000-4250</td>
</tr>
<tr>
<td></td>
<td>2750-3600</td>
</tr>
<tr>
<td></td>
<td>1500-2800</td>
</tr>
<tr>
<td>$D_3$ (X – 2X)</td>
<td>2000-3820</td>
</tr>
<tr>
<td></td>
<td>1800-2450</td>
</tr>
<tr>
<td></td>
<td>1300-1400</td>
</tr>
<tr>
<td>Medium-action video</td>
<td></td>
</tr>
<tr>
<td>$D_1$</td>
<td>3270-7350</td>
</tr>
<tr>
<td></td>
<td>2500-6500</td>
</tr>
<tr>
<td></td>
<td>2500-5300</td>
</tr>
<tr>
<td>$D_2$</td>
<td>2920-3900</td>
</tr>
<tr>
<td></td>
<td>1900-3000</td>
</tr>
<tr>
<td></td>
<td>1100-2750</td>
</tr>
<tr>
<td>$D_3$</td>
<td>1950-2650</td>
</tr>
<tr>
<td></td>
<td>1500-1590</td>
</tr>
<tr>
<td></td>
<td>950-1300</td>
</tr>
<tr>
<td>Low-action video</td>
<td></td>
</tr>
<tr>
<td>$D_1$</td>
<td>3000-7150</td>
</tr>
<tr>
<td></td>
<td>2500-6000</td>
</tr>
<tr>
<td></td>
<td>1100-4950</td>
</tr>
<tr>
<td>$D_2$</td>
<td>1650-3700</td>
</tr>
<tr>
<td></td>
<td>1300-2900</td>
</tr>
<tr>
<td></td>
<td>950-2700</td>
</tr>
<tr>
<td>$D_3$</td>
<td>1450-1900</td>
</tr>
<tr>
<td></td>
<td>900-1500</td>
</tr>
<tr>
<td></td>
<td>600-1200</td>
</tr>
</tbody>
</table>
allocation to 6 users (2 in each priority level high, medium and low) with and without applying the feedback mechanism. It can be seen that, the clients with the feedback enabled, receive a varying bitrate with respect to time, while the clients who do not send any feedback receive the same bitrate for the entire transmission.

According to the feedback received, there is a gradual reduction in the quality obtained by clients in medium and low priority, and a subsequent increase in that of the client having the highest priority simultaneously. In each of the envisaged scenarios 95-99% of the total available bandwidth is used for video and multimedia communication and the remaining 1-5% is reserved for feedback purpose. Fig. 5(a) and 5(b) compare the resource allocation without using PRiMA and using PRiMA respectively. It is observed that without using PRiMA all clients, regardless of their device-resolutions, receive same bitrate and hence the same QoS, whereas when PRiMA is applied, the same device types receive different bitrates (and therefore different QoS) according to the services chosen.

Apart from the bandwidth and the number of clients and their priorities, the bitrate depends on the number of bits/symbol, \( n \). In case of QPSK modulation technique, \( n = 2 \), whereas in case of 8-PSK modulation technique, \( n = 3 \). A higher modulation technique requires higher SINR (signal to interference noise ratio) at the receiver of the communicating link, but at the same time would result in higher \( Q \). In this simulator, the video streams are assumed to be modulated with QPSK modulation technique, i.e., \( n = 2 \). The perceived quality of a user is as given by (1) where \( r \) is the bitrate received by the client.

In order to assess \( Q \) as in (1), different constants are used for different video types, according to their action content [25]. For example \( Q_0 = 5.225, \chi_0 = -0.045, \chi_k = 124.762, \zeta_k = 1.116 \) and \( \chi_k = -33.9 \) are used for the case of high-action video transmissions. For low-action video sequences the values for \( Q_0 \) and \( \chi_0 \) would be given by, \( Q_0 = 5.062, \chi_0 = -0.025 \), with other constants being the same. For medium-action video transmissions, constants have been calculated based on the average values \( Q_0 = 5.115, \chi_0 = -0.035 \); the other constants are kept the same. The bitrate \( r \), as shown in (2), is directly proportional to bandwidth of a particular priority-holding client, to be precise, the bandwidth is half of the bitrate, as QPSK technique is used here.

In the proposed topology, the client requests the video and chooses the type of service and the category in which resolution of its device belongs. Then taking into account all these priorities and assuming other parameters like rate and bit depth to be constant, the bit-rate is set for the client (as explained in Section IV. B).

C. Relation between Resolution and Bitrate with respect to Perceived Quality

The relation between resolution of screen and minimum bitrate required in order to obtain the desired perceived quality is specified in Fig. 6 [29]. An MPEG bitrate is not a linear function. As shown in the graph, after a certain range the upper limit will reach a plateau where adding more bitrate will not make the quality any better. The graph in Fig. 6 shows
Fig. 5(a): QoS observed without prioritized resource allocation

Fig. 5(b): QoS observed with prioritized resource allocation using PRiMA

(Advance online publication: 13 May 2010)
the ideal received quality ranges based on the allocated bitrate and the resolution of devices. The quality is a relative quality for certain resolution as medium resolution device at maximum bitrate will have less detail than high resolution device at its respective maximum bitrate.

VII. RESULTS

The results shown in the Tables III, IV and V have been formulated for medium-action video. Similar simulations can be done for low and high action videos as well. The case when all the clients have same \((C, D)\) pair, for instance, all 6 clients opt for highest client priority and are using PDAs as device, therefore having same priority pair as \((1, 2)\). In this case all the clients get equal bandwidth of \(B/6\) and uniform \(Q\)-value of 3.44 \((Q_{avg})\).

Table III is for a case where all clients have same \(C\) and different \(D\). As for instance, all the users have client priority \(C\) but possess different devices. Hence the priority pair becomes \((C, i)\), where \(i = 1, 2, 3\). Table IV is similar to table III, but with same device priority \(D\) and different client priorities, therefore the priority pair is \((i, D)\).

Table V shows the simulation results for medium action video content, where the clients are characterized by their device types \((D)\) and service chosen \((C)\).

It can be seen that clients with same \(C\) but different \(D\) will receive different bitrates and hence different QoS, since the bitrates are now allocated incorporating the effect of device heterogeneity and also with respect to the other simultaneous users. Table V shows the respective perceived quality values as obtained by different clients with different priorities for service and devices chosen in the heterogeneous network.

It can be concluded from Table V that majority of clients receive QoS which is more than \(Q_{avg}\). Also, it is shown that the clients that receive QoS value less than 2 belong to the least priority group both in terms of the service chosen and device type. Whereas other clients with higher priorities receive QoS values around 4 (significantly greater than \(Q_{avg}\)) and hence better video quality, which was not possible without prioritization when unbiased distribution of network resources was done. ‘x’ in the table denotes that no

**Table III:** Perceived Quality by clients with same \(C\) and different \(D\).

<table>
<thead>
<tr>
<th>No. Of users</th>
<th>((C, D))</th>
<th>Perceived Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>((i, 1))</td>
<td>3.923</td>
</tr>
<tr>
<td>2</td>
<td>((i, 2))</td>
<td>3.347</td>
</tr>
<tr>
<td>2</td>
<td>((i, 3))</td>
<td>2.536</td>
</tr>
</tbody>
</table>

**Table IV:** Perceived Quality by clients with same \(D\) and different \(C\).

<table>
<thead>
<tr>
<th>No. Of users</th>
<th>((C, D))</th>
<th>Perceived Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>((1, i))</td>
<td>3.907</td>
</tr>
<tr>
<td>2</td>
<td>((2, i))</td>
<td>3.457</td>
</tr>
<tr>
<td>2</td>
<td>((3, i))</td>
<td>2.196</td>
</tr>
</tbody>
</table>

(Advance online publication: 13 May 2010)
client with the particular combination of client and device priorities, \((C, D)\) is present. The focus of the results shown in Fig. 7 is on the sum of client and device priorities, i.e. \(C + D\). The motivation for choosing the sum of priorities for analysing the received QoS is to analyze the benefits of PRiMA in applying prioritization in the network. By using the sum of priorities we can analyze the importance of applying prioritization in order to effectively capture the network heterogeneity. When the type of video to be streamed (high, medium or low) is fixed, the bitrate to be allocated then depends upon the two other priorities i.e. service chosen \((C)\) and device type \((D)\). In case of same priority sum i.e. sum of the service priority \(C\) and device priority \(D\), different possible combinations of \(C\) and \(D\) with the given constraints viz. \(1 \leq C, D \leq 3\); have been considered and their average has been plotted. Hence, for different combinations of \(C\) and \(D\), sum of priorities of all the clients in the network have been considered, which can range from 12-36. The sum is 12 when all clients have priority pair \((1, 1)\), and 36 when all clients have \((3, 3)\). In both these cases and with sum = 24, the bitrate allocation would be equal as all the clients would possess similar devices, so no prioritization can be incorporated. Hence, the Fig. 7 shows values of sum in range 12 ≤ 32. It shows that the average perceived quality of the client with \(C\) as 1 or 2 is

### Table V: Perceived Quality by clients with different combinations of \(C\) and \(D\).

<table>
<thead>
<tr>
<th>No. of Users with respective priorities</th>
<th>Respective Perceived Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(C1)</td>
</tr>
<tr>
<td></td>
<td>(D1)</td>
</tr>
<tr>
<td>Clients with unique ((C, D))</td>
<td></td>
</tr>
<tr>
<td>0 1 1 0 1 1 1 1 1 0 x</td>
<td>4.10</td>
</tr>
<tr>
<td>1 0 1 1 1 0 1 1 1 1 1 0 1 1 1 1 1 x</td>
<td>3.73</td>
</tr>
<tr>
<td>0 1 1 1 0 1 1 1 0 1 0 x</td>
<td>3.65</td>
</tr>
<tr>
<td>0 1 1 1 1 0 1 0 1 x</td>
<td>4.01</td>
</tr>
<tr>
<td>More than 1 client with same device and client priority</td>
<td></td>
</tr>
<tr>
<td>1 0 1 0 1 1 0 1 1 0 1 0 0 x</td>
<td>4.25</td>
</tr>
<tr>
<td>0 2 0 2 0 0 0 0 1 0 0 0 0 x</td>
<td>3.98</td>
</tr>
<tr>
<td>1 1 0 1 0 0 1 0 1 0 0 1 1 1 0 0 x</td>
<td>4.12</td>
</tr>
<tr>
<td>0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 x</td>
<td>x</td>
</tr>
<tr>
<td>0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 x</td>
<td>3.96</td>
</tr>
<tr>
<td>1 1 0 2 0 0 0 1 0 1 1 1 1 0 1 0 1 x</td>
<td>4.07</td>
</tr>
<tr>
<td>1 2 0 0 0 0 1 0 1 0 0 1 0 1 x</td>
<td>3.99</td>
</tr>
</tbody>
</table>

**Fig. 7:** Average perceived quality of clients in the highest client priorities \((C1\) and \(C2\)), and Q-avg. without prioritization relative to priority sum \((C + D)\)

(Advance online publication: 13 May 2010)
considerably above the average Q-value received without prioritization. This is an extremely significant result as it shows that the average perceived quality of clients with higher priorities is significantly more than the average quality without prioritization.

Fig. 7 also shows that as more and more clients chose different priorities for the service demanded (C) and the end-user device (D), i.e. as the network heterogeneity increases, significant improvement is observed in the QoS received by majority of clients by applying PRiMA. This signifies the importance of prioritization in improving the end-user quality for prioritized users given the network resources. Also, it can be concluded that as the heterogeneity in the network increases (sum of C and D increases); more freedom is available with the service providers to distribute the available bandwidth and hence incorporate the proposed prioritization efficiently.

Hence, we can see the importance of PRiMA in improving the QoS of majority of clients in the network by considering the combined effects of client and device priorities for a given video content. This shows that PRiMA efficiently uses prioritization and effectively captures the heterogeneity in the network.

VIII. CONCLUSION AND FUTURE WORK
This paper presents a novel prioritization mechanism for adaptive multimedia transmission in two-hop heterogeneous networks. The prioritization mechanism proposed not only incorporates device resolution, but also increases the perceived quality for most categories of users. A novel feedback mechanism is also proposed and implemented which distributes the bandwidth with respect to both client and the device priorities. Results show that the client priorities of users of categories 1 and 2 are much above the average Q received as shown in Table III. These results signify the importance of PRiMA in providing a real time dynamic adaptive mechanism that not only improves the EUPQ but also provide better and optimum service to different users. Also, the prioritization ensures that the available network bandwidth is utilized effectively and the priorities chosen by the user are directly reflected in allocation of the network resources to different users.

In this paper we have shown that the average perceived quality of clients with higher priorities is significantly more than the average quality without prioritization. This signifies the importance of prioritization in improving the end-user quality for prioritized users given the bandwidth constraints. An important task for future research work is to theoretically analyze the prioritization aspect in PRiMA and find out the optimal values for the bandwidth allotment ratios that would maximize the perceived quality of each category and of the entire network. As a part of future work PRiMA can be tested using WiMax Access Point (WAP-400) [27].

Another factor that will affect the perceived quality of an end-user is mobility. Though, in this paper, the performance of PRiMA was analyzed assuming the end-user to be stationary, future work can also involve client mobility (with variable speed/direction). In combining the Proxy Client and Proxy Server model, the feedback-based QOAS scheme can be applied between both BS and proxy node (PC/PS) and between PC/PS and the end-user. This would increase the overall QoS of the end-to-end link. However, that would increase the overall complexity of the system and a potential increase in the delivery time. Hence future work can focus on this dual feedback approach with an intelligent proxy node.

ACKNOWLEDGMENT
The authors would like to thank Science Foundation Ireland (SFI) for the Online Dublin Computer Science Summer School (ODCSSS) program and the Irish Research Council for Science Engineering and Technology (IRCSET) for their support and encouragement.

REFERENCES


