A Mobile QoE Architecture for Heterogeneous Multimedia Wireless Networks

Carlos Quadros, Eduardo Cerqueira Federal University of Para Belem, Brazil {jean, cerqueira}@ufpa.br Augusto Neto Federal University of Ceara, Brazil Institute of Telecommunications, Aveiro, Portugal augusto@deti.ufc.br Andre Riker, Roger Immich, Marilia Curado University of Coimbra Coimbra, Portugal {riker, immich,marilia}@dei.uc.pt

Abstract— One of the main requirements in this emerging wireless multimedia era is the Quality of Experience (QoE) assurance for 2D or 3D video applications in heterogeneous multi-operator environments. Therefore, this paper proposes a QoE Architecture for Heterogeneous Multimedia Wireless Networks, called QoEHand. QoEHand extends the Media Independent Handover (MIH)/IEEE 802.21 proposal with QoEawareness, seamless mobility, dynamic class of service mapping and a set of content adaptation schemes. The proposed solution allows the best connection and considers the QoE needs of mobile clients and available wireless resources in IEEE 802.11e and IEEE 802.16e service classes. Simulation experiments were carried out to show the impact and benefits of QoEHand on the user's perception, by using objective and subjective QoE metrics.

Keywords: Multimedia; MIH; QoE; heterogeneous networks

I. INTRODUCTION

The integration of heterogeneous wireless networks, such as IEEE 802.11 and IEEE 802.16, Long Term Evolution (LTE), in multi-access and multi-operator systems is bringing revolutionary changes to the Internet by introducing new opportunities, better communication channels and the possibility of providing better Quality of Experience (QoE) assurance for users of wireless services. The multi-operator wireless environment will allow mobile users to be "Always Best Connected" (ABC) (including during handover), where seamless mobility will be coupled with respect to user's preferences. In this scenario, the creation of novel architectures is required to allow vertical/horizontal seamless QoE-aware handovers in heterogeneous wireless networks.

As a means of providing interoperability and seamless mobility in heterogeneous systems, the IEEE introduced the standard called IEEE 802.21 or MIH (Media Independent Handover Services) [1]. The MIH is a middleware for heterogeneous networks, where it has a set of protocols and mechanisms to allow different IEEE or non-IEEE wireless technologies to be integrated, while ensuring both vertical and horizontal handovers. However, MIH alone is not able to provide either an ABC approach or QoE assurance of media applications over wireless clients.

In traditional handover schemes, such as MIH, users are connected to access points that offer the best power present in a received radio signal or Quality of Service (QoS) metrics, such as loss, delay and jitter values. In heterogeneous networks, the Received Signal Strength Indicator (RSSI) and QoS metrics by themselves, are not sufficient to support QoE-aware seamless mobility. This is because that metrics cannot capture the subjective aspects of multimedia content with regard to user perception/satisfaction [1][2][3]. In view of this, heterogeneous wireless QoE architectures with quality estimator, mapping and adaptation support are recognized as key issues for the success of future systems [4].

A seamless mobile multimedia architecture with a video quality estimator scheme, is designed aiming at estimating user perception in the currently connected and candidate access points (to be will be used in the handover decision process) and must integrate comprehensive monitoring schemes with service metrics such as visual codec, Group of Pictures (GoP) length, intra-frame dependency, spatio-temporal video activity, network impairments and other relevant factors, such as the capacity of wireless systems/service classes [5][6]. QoE-aware prediction models can be devised through efficient, network cross-layer agnostic content-awareness, QoE monitoring and Artificial Intelligence (AI) techniques along with the corresponding cognitive evaluation of user's inputs [7].

One of the key issues when deploying heterogeneous wireless systems is that each domain may support different QoS models (e.g., IEEE 802.11e or IEEE 801.16), offer the same wireless service classes with different definitions or even service classes with different compositions [8]. Therefore, a QoE mapping mechanisms must be used to map application requirements and user's perception into available wireless service classes, on the basis of information about the available service classes within or between wireless networks (multiaccess and multi-operators) and the video quality level score given by the quality estimator. During periods of congestion in a selected service class, the adaptation mechanism must adjust the quality level of the 2D/3D multimedia applications, by selecting another service class to map on-going packets or by dropping packets in overloaded queues according to their impact on user's perception.

This paper presents a mobile QoE Architecture for Heterogeneous Wireless Multimedia Networks (QoEHand). QoEHand extends MIH with QoE-aware video quality estimator, mapping and adaptation control. Simulation experiments were carried out to measure the impact and benefits of the proposed solution in a multi-operator IEEE 802.11e and IEEE 802.16e system on the user's perception, by employing objective and subjective QoE metrics.

This paper is structured as follows. The related works are discussed in Section 2. Section 3 details the QoEHand.

Performance evaluation results are analysed in Section 4. Section 5 summarizes the paper and presents the future works

II. RELATED WORKS

A video quality estimator proposal that relies on the structure (width, length and height) of the video to provide a QoE quality score is introduced by [9]. However, this proposal was not evaluated in wireless networks and does not take video motion and complexity levels and intra-frame dependency into account in the prediction process.

Another video quality estimator, known as Pseudo-Subjective Quality Assessment (PSQA) [10] and, its extensions [11], uses a Random Neural Network to map network impairments and video characteristics of the user's perception. Different applications have already used PSQA, such as in the context of video streaming over Peer-to-Peer networks [12], voice [13], VoIP [14], video streaming over WLAN [15] and a scheme for network selection together with an admission control mechanism for IEEE 802.11 [16]. However, PSQAbased solutions only use QoS parameters, such as packet loss rate as input and do not consider videos for the prediction models that have different motion and complexity levels.

The main challenges and functionalities required to create an IEEE 802.21 Media Independence Service Layer to optimize the usage of resources in heterogeneous wireless networks is discussed in [17], by adopting a modular and selforganized approach. Control modules, such as mobility, routing and video quality estimator, can be easily integrated into the architecture. Our proposal uses the same modular and selforganized approach, but also includes and analyses the benefits of a MIH system integrated with QoE-awareness video quality estimator, dynamic mapping and a set of adaptation schemes.

An enhanced information server for seamless vertical handover in IEEE 802.21 MIH networks is proposed in [18]. Information about the wireless channel conditions is assessed and used to provide seamless mobility. However, this proposal does not assume the existence of networks with different class of services as expected in future wireless systems. It also lacks QoE assessment and optimization support. The IEEE 802.21-based solution [19] draws attention to the benefits of its proposal by analyzing packet loss, handover latency and handover discovery time. Our proposal follows the same make-before-break approach to provide seamless handover, but it also introduces the QoE video quality estimator, mapping and adaptation support as required for heterogeneous networks.

A QoS architecture to provide a level of quality assurance for applications in heterogeneous environments is presented in [20]. The proposal uses a schedule-based scheme that draws on information about delay, loss and current network resources, and adjusts the scheduler to improve the video quality of delivery. However, this work does not provide seamless handover control or follow its procedures in accordance with the user's experience/QoE scheme (only QoS parameters).

An environment-aware interface management framework for QoS support is used in multi-interface wireless terminals (IEEE 802.11e and IEEE 802.16e) [21]. A static mapping approach between service classes of different QoS models is included in the framework to provide seamless mobility. However, in multi-operator wireless systems, a static mapping model is not suitable because the heterogeneous networks might offer the same wireless service classes but with different definitions or even service classes with different compositions. In addition, this QoE-unaware proposal does not support adaptation mechanisms for congestion periods.

The main challenges for optimizing QoE in Next Generation Networks (NGN) are discussed in [22]. However, our solution will succeed in implementing and validating a QoE handover architecture for converged next generation wireless heterogeneous networks, by extending MIH and its traffic/mobility controllers with QoE video quality prediction, mapping and adaptation schemes.

III. MOBILE QOE ARCHITECTURE FOR HETEROGENEOUS WIRELESS MULTIMEDIA NETWORKS (QOEHAND)

QoEHand follows a self-adaptation principle and has been devised on the basis of a modular integration of control components over application, session and network layers (see Figure 1). The self-adaptation principle assures QoEHand to adjust its control functions to different wireless infrastructures and service classes. The modular design allows the inclusion (or change) of policies, technologies and emerging services.



Fig 1. QoEHand mechanisms, protocols and interfaces

The QoE mapping maps application requirements and user's perception into available wireless service classes (IEEE 802.11e or IEEE 801.16 QoS models). The mapping process is performed by drawing on information about the available service classes within or between networks (in multiple paths when possible), application QoS/QoE requirements, the video quality estimator score and mapping policies. The latter decides which, what and when mapping methods must be used to carry out a request. After the mapping decision, the QoS scheduler is triggered to map the packets in the selected service class.

The mapping policies define two main mapping methods to select the best class for an emerging multimedia application (its flows/components), called Full and Partial-Matching. A fullmatching mapping is accomplished when the quality level score of an application in a class is better than the minimal level. If there is more than one class result in the same quality level score, the policy scheme only considers the service class, which has more available resources in terms of bandwidth. If the most suitable wireless service class is unable to assure a full matching (due to congestions or service classes with different configurations in terms of loss, delay and jitter support), the adaptation scheme is triggered to seek a potential adaptation of the applications that can fit the current network conditions.

Depending on the business strategies, the nature of the multimedia content and the video quality level score, a set of dynamic partial matching mapping approaches can be applied as follows: Downgrade class mapping (i): This approach chooses a less important class to accommodate the application that assures a good/acceptable level of quality (video quality estimator score >= minimal video quality level requirement) (ii): This approach selects service classes according to the order of priority of different multimedia components. Video communication is much more sensitive to packet loss than audio communication, because the human eye can often detect small glitches in a video stream caused by relatively minor packet loss, to an extent whereby enjoyment and/or understanding are more severely affected. For example, since voice has a higher priority than visual content in an application, the packets of audio flows are mapped to the best class and the packets of video flows to a lower priority class; Hierarchical 2D/3D mapping (iii): This scheme maps 2D or 3D video frames based on the importance of each frame type from the user's experience, such as, packets carrying I frames are allocated to the best class, while packets of P and B frames are mapped in a less important class.

An advanced video quality estimator mechanism is implemented by QoEHand agents as detailed in [5]. This scheme allows real-time video quality prediction in multipath and multi-operator heterogeneous networks, where it gives a video quality score which corresponds as closely as possible to the user's perception. This value is used for monitoring the quality level of the multimedia content consumed by a wireless user, as well as to predict the user's experience of video streaming in multipath and multi-operator systems (used as input for mapping and adaptation procedures). This mechanism is able to carry out its video quality prediction on the basis of the statistical learning of Neural Networks (NN) because of its ability to learn and adapt from different environments. During the training phase, the learning process is iterative, and the NN improves its performance gradually as it interacts with the environment, and in this way, the process results in a better generalization. After the training phase, the proposed video quality estimator is used in real-time as a function that maps the input parameters for the MOS.

The video quality estimator uses objective parameters from the video encoder and wireless network conditions, as well as information about the perception of real viewers collected from the MOS experiments. The MOS is the most widely used subjective metric and is recommended by the ITU [26]. The MOS is obtained by asking people to grade the quality of video on a five-point scale (Excellent, Good, Fair, Poor and Bad).

The video quality estimator takes into account the current network conditions (different network impairments in the service classes) and different video parameters (percentage of losses in I, P and B frames, total number of losses, GoP length, and motion and complexity levels) that directly affect the quality of the video in terms of MOS prediction. As a result, this procedure gives a video quality score that corresponds as closely as possible to human perception in real-time. Since our proposal has been tested and validated, it is a dynamic and content-aware quality predictor that is able to estimate the video quality of several types of video content features in realistic multi-operator converged network conditions, without any interaction with real viewers. Figure 2 shows an overview of the video quality estimator mechanism, where original videos are encoded, transmitted in wireless networking systems (suffered network impairments) and the quality level of received videos are subjectively evaluated by real observers.



Fig 2. QoE video quality estimator

This paper validated the video quality estimator in an IEEE 802.11e and IEEE 802.16e system as explained in Section 4. After observers have evaluated each video in service classes with different congestion levels, the training process was performed by the training video database to obtain the mapping between the selected input (video/network) parameters and the MOS. The validation task was done with cross-validation techniques to minimize the generalization error. If the predicted MOS indicates a low level of quality for the video flow, the QoEHand will search a new (most suitable) class to map or adapt the multimedia content. The QoEHand considers the service classes available in the current and target networks so that it can offer continuous and seamless services and a satisfactory multimedia content delivery in heterogeneous networks, as expected in future environments.

As mentioned earlier, one problem arising from multioperator wireless systems is the fact that each network provider can support different QoS models (e.g., IEEE 802.11e or IEEE 801.16) and can offer the same class of service with different definitions. For this reason, when the mapping process is not optimal (perfect match), the QoE Adaptation adjusts (e.g., downgrades) the quality level of the emerging applications if the network resources in a service class are unavailable (e.g., in congestion periods). The downgrade adaptation process is reversible when there are available resources in the previous service class again. In this case, the QoEHand can trigger MIH to handover the wireless client to the old network and maps all the flows into the previous service class. Since the success of our seamless proposal depends on adopting a make-beforebreak approach, the resources that are allocated and not used in previous or candidate the service classes, are released by softstate operations, for instance after a handover.

A set of network adaptation profiles can be performed by the adaptation mechanism to control the quality level of new or current applications as follows: 2D/3D Frame dropping adaptation (i): This approach drops packets according to the visual importance of each frame encoded with common hierarchical 2D and 3D MPEG/H.264 codecs. I frames are marked with low priority and B frames with high priority on dropping probability. Due to the intra-fame dependency in hierarchical codecs, when a P-frame is discarded, all of the subsequent P frames and B frames within the same GoP should also be discarded. When an I frame is discarded, all other frames within the same GoP is dropped. Scalable video adaptation (ii): This approach adjusts the quality level of applications by, dropping or adding low important flows of scalable multimedia data. Hierarchical component adaptation (iii): Media flows within an application can be marked with different priorities. Audio packets are marked with low priority and video with high priority dropping probability.

After introducing the functionalities of the mapping, adaptation, and video quality estimator, QoEHand will be described and integrated with an IEEE 802.21 converged network. In this system, each BS or AP (both with MIH) gives information before the connection, about which service class (including the current channel conditions in terms of loss, delay, jitter and bandwidth) are available to map the applications of the wireless clients. After selecting a service class (e.g., Access Category 1 in IEEE 802.11e or Unsolicited Grant Service (UGS) in IEEE 802.16e), the BS/AP establishes a connection linked to the user in a QoS class with enough resources to ensure a quality level for the application.

When a wireless node detects a Target/Candidate/Foreigner Network, the MIH module sends a MIH LINK DETECTED message to inform the Target Network that there is a new client in a coverage area. The Target Network sends a MIH_LINK_PARAMETERS REPORT to the Current Network, where the available service classes and its conditions are contained. After this, the QoEHand in the Current Network measures and compares the level of video quality in the current and foreign service classes and triggers the mapping mechanism to select the best service class for the multimedia applications. If the quality level score in the target network is higher, QoEHand informs the mobile node of the handover decision by sending a MIH_LINK_PARAMETERS_REPORT. The seamless handover is initiated by using а MIH HANDOVER INITIATE. The handover can also be triggered in congestion periods when the video quality estimator detects a service class in a target network that will assure a better level of quality. The MIH establishes communication between the lower and upper layers, based on a set of primitives defined as SAPs (Service Access Points) and presented in its standard.

IV. PERFORMANCE EVALUATION AND RESULTS

Performance evaluations were carried out by Network Simulator 2 (NS2) and Evalvid Tool. The objective was to analyse the benefits of QoEHand in converged IEEE 802.11e and IEEE 802.16e networks, compared with a system without QoEHand (without video prediction, mapping and adaptation – only MIH functionalities), by measuring objective and subjective metrics. The QoEHand Quality Estimator Neutral Network was formed with the aid of MATLAB. The Real Time Protocol (RTP) payload header includes a field which indicates the current frame type, i.e., I, P, or B frames.

Four profiles were configured in the system to determine the benefits of QoEHand with different scenarios and experiments: (i) Pure_MIH (without QoEHand); (ii) QoEHand_Full, when a full mapping match is achieved during the handover and there are available resources in the service class of the foreign network; (iii) QoEHand_Part profile, which re-maps all the packets of a video sequence into a less important class in the target network, because the most suitable wireless service class cannot assure a full-matching (e.g., due to congestions); (iv) QoEHand_Drop which controls the video quality level by dropping video packets in descending order of importance from the standpoint of the user's perception.

The Structural Similarity Metric (SSIM) QoE metric is based on frame-to-frame measuring of three components (luminance similarity, contrast similarity and structural similarity) and combining them into a single value, called the index. The SSIM index is a value between 0 and 1, where 0 stands for zero correlation with the original image, and 1 stands for exactly the same image.

The ITU-T MOS recommendation was used for subjective evaluation with 55 observers. They had normal vision and their ages ranged from 18 to 45. The observers included undergraduate students, postgraduate students, and university staff. The test platform used was a Desktop PC with Intel Core i5, 4GB RAM and a 21" LCD monitor. 10 different wellknown Internet video sequences were selected for the experiments (Akiyo, Container, Coastguard, Highway, Football, Hall, Mobile, Grandma, News, Silent) with different complexity and motion levels [23]. The video sequences were encoded in MPEG4 format and the duration varied from 10s to 30s. The GoP length was 18, which is what can be expected for common Internet videos. To provide a large enough video database and increase the system reliability, each selected video was simulated 10 times by varying the congestion periods (from 0% to 50% in steps of 5) in a class, resulting in a total of 100 (received) videos with different packet loss rate.

The multi-operator use case scenario is composed of IEEE 802.11e and 802.16 networks, where QoEHand seeks to control the quality level of the multimedia applications and provide seamless mobility and adaptation in congestion periods. Two service classes are configured in each network (IEEE 802.11e - AC_V0 and AC_V1 / IEEE 802.16e - Realtime Polling Service (rtPS) and Non-real-time Polling Service (nrPS). The service class where the user is receiving the video in the current network will experience congestions from 5% to 50% in steps of 5 caused by concurrent traffic. Hence, QoEHand will interact with MIH to adjust (handover, remapping or drop packets) the video quality level based on one of its 3 profiles. 10 simulations were carried out for each video, where, in five of them, the receiver experienced congestions in a service class of an IEEE 802.11e network. In the other 5 experiments, the receiver experienced congestions in a service class of an IEEE 802.16e network.

The MOS results (Figure 3) show that QoEHand assures an excellent quality level for the videos during congestion periods when the QoEHand_Full and QoEHand_Part profiles are configured in the system. In the QoEHand_Part, the videos still have a good-to-excellent quality level even when re-mapped to a less important class (nrPS or AC_V1) with a packet-loss rate of approximately 2%. The QoEHand_Drop attempts to keep

the application at an excellent quality level of up to 10% of congestion and at a good/regular quality level of up to 30% of congestion. When the Pure_MIH is used, the video quality level was considered poor by all the observers if there was at least 10% of congestion in a wireless service class.



Fig 3. Congestion x MOS for all the profiles

Figure 4 illustrates the SSIM values obtained from all the experiments. When the QoEHand_Full is used, the SSIM value is, on average, approximately 7%, 21% and 40% better compared with the profiles of QoEHand_Part, QoEHand_Drop and Pure_MIH, respectively.



Fig 4. Congestion x SSIM for all the profiles

V. CONCLUSION AND FUTURE WORKS

Future multimedia systems envision the synergy among heterogeneous wireless communications will co-exist, where users must be connected to networks able to provide the best QoE for his/her applications. This paper proposes QoEHand to improve the user's perception and optimize the usage of wireless network resources in a competitive converged system. QoEHand extends MIH with QoE assurance by coordinating quality estimator, mapping and adaptation mechanisms. Due to its modular approach, QoEHand can be adjusted to operate with different wireless technologies, such as LTE.

Performance evaluations were carried out to present the impact and benefit of QoEHand in an IEEE 802.11e and IEEE 802.16e multi-operator system. The results obtained show that QoEHand provides a better quality level for multimedia applications compared to a pure MIH scheme. For instance, QoEHand_Full and QoEHand_Part profiles keep the video with an excellent MOS during all experiments.

ACKNOWLEDGEMENT

This work was funded by Coordination for the Improvement of Higher Level Personnel (CAPES), National Council for Scientific and Technological Development (CNPq), Fundação Amazônia Paraense de Amparo à Pesquisa (FAPESPA) and FCT UBIQUIMESH project (PTDC/EEATEL/105472/2008).

REFERENCES

- M. Venkataraman and M. Chatterjee, "Inferring video QoE in real time", Network IEEE, vol. 25, no. 1, pp. 4–13, January-February 2011.
- [2] Jingjing Zhang; Ansari, N.; "On Assuring End-to-End QoE in Next Generation Networks: Challenges and a possible Solution", Communications Magazine IEEE, vol 49, no 7, pp. 185-191, July 2011.
- [3] R. Sousa *et all*, "GoP size influence in high resolution video streaming over wireless mesh network", in Computers and Communications (ISCC), 2010 IEEE Symposium on, Riccione, Italy, 2010.
- [4] Huszak, A.; Imre, S., "Analysing GoP structure and packet loss effects on error propagation in mpeg-4 video streams", in Communications, Control and Signal Processing, 4th Int. Symp. on. IEEE, March 2010.
- [5] Aguiar, E at all, "Real-Time QoE Prediction for Multimedia Applications in Wireless Mesh Network", 4th IEEE Future Multimedia Network Workshop, Las Vegas, USA, Jan 2012.
- [6] Keimel, C.; Rothbucher, M.; HaoShen; Diepold, K.; "Video is a cube". IEEE Signal Processing Magazine, Vol 28, Issue 6, November 2011.
- [7] S. Mohamed and G. Rubino, "A study of real-time packet video quality using random neural networks", Circuits and Systems for Video Tech., IEEE Transactions on, vol.12, no.12, pp.1071–1083, December 2002.
- [8] M. Ghareeb and C. Viho, "Hybrid QpE assessment is well-suited for multiple description coding video streaming in overlay networks", in 8th Com. Networks and Services Research Conf. IEEE, 2010, pp.327–333.
- [9] H. Cancel et all, "A robust p2p streaming architecture and its application to a high quality live-video service," Electronic Notes in Discrete Mathematics, vol. 30, pp. 219–224, 2008.
- [10] A. da Silva, M. Varela, E. de Souza e Silva, R. Leao, and G. Rubino, "Quality assessment of interactive voice applications," Computer Networks, vol. 52, no. 6, pp. 1179–1192, 2008.
- [11] G. Rubino, M. Varela, and J. Bonnin, "Controlling multimedia qos in the future home network using the psqa metric," The Computer Journal, vol. 49, no. 2, p. 137, 2006.
- [12] K. Piamrat, C. Viho, J. Bonnin, and A. Ksentini, "Quality of experience measurements for video streaming over wireless networks", IEEE Int. Conf. on New Generations in Inf. Tech., Las Vegas, USA, April 2009.
- [13] K. Piamrat, A. Ksentini, C. Viho, and J. Bonnin, "Qoe-based network selection for multimedia users in ieee 802.11 wireless networks," in Local Computer Networks. Conference on. IEEE, 2008, pp. 388–394.
- [14] De La Oliva et all,"An overview of IEEE 802.21: media-independent handover services", IEEE Wireless Communications, V15, August 2008.
- [15] Kim, Younghyun *et all*, "An enhanced information server for seamless vertical handover in IEEE 802.21 MIH networks", IEEE Computer Network, vol 55, Issue 1, January 2011.
- [16] Lim, Wan-Seon *et all*, "Implementation and performance study of IEEE 802.21 in integrated IEEE 802.11/802.16e networks", Computer Communications, Vol 32, Issue 1, January 2009.
- [17] Liang Zhou et all, "Joint Forensics-Scheduling Strategy for Delay-Sensitive Multimedia Applications over Heterogeneous Networks", IEEE Journal on Selected Areas in Com., Vol 29, Issue 7, August 2011.
- [18] HE, Q., "A Fuzzy Logic Based Vertical Handoff Decision Algorithm between WWAN and WLAN", IEEE International Conference on Networking and Digital Society, Shanghai, China, May 2010.
- [19] Andi, C. *et all*, "Environment-Aware QoS framework for multi-interface terminal", Computer Communications, Vol. 33, Issue 9, June 2010.
- [20] Jingjing Zhang; Ansari, N., "On assuring end-to-end QoE in next generation networks: challenges and a possible solution", IEEE CommunicationsMagazine. Vol. 49, Issue: 7, 2011.
- [21] Chakchai So-In; Jain, R.; Tamimi, A.-K., "Scheduling in IEEE 802.16e mobile WiMAX networks: key issues and a survey", IEEE Journal on Selected Areas in Communications, Volume: 27, 2009.
- [22] Recommendation ITU-T BT. 500-11. accessed in 2012.
- [23] "Video trace library," Available in trace.eas.asu.edu, accessed in 2012.