Feedback-free Early VoIP Quality Adaptation Scheme in Next Generation Networks

Is-Haka Mkwawa, Emmanuel Jammeh, Lingfen Sun and Emmanuel Ifeachor
School of Computing and Mathematics
University of Plymouth, Plymouth, PL4 8AA, UK
(Is-Haka.Mkwawa,Emmanuel.Jammeh, L.Sun,e.ifeachor)@plymouth.ac.uk

Abstract—The Next Generation Networks (NGNs) were proposed to provide ubiquitous access to multimedia services over the Internet. With the best effort nature of the Internet, the quality of real time multimedia services such as voice and video can rapidly degrade due to network impairments.

In this paper, we propose a novel, feedback-free Early VoIP Quality Adaptation Scheme (EVQAS) which can adapt voice send bit rate (e.g. Adaptive Multi-Rate (AMR) codec mode) automatically according to the buildup of queuing delay in the network. This differs from existing adaptation methods which adapt send bit rate according to either packet loss information or Perceived Quality of Service (PQoS) metrics (e.g. Mean Opinion Score (MOS)) derived from packet loss and delay. The proposed scheme can avoid congested packet loss which has an obvious impact on perceived voice quality. We implemented this adaptation mechanism in open Android mobile and tested in Open IMS Core testbed. Preliminary results show that the early adaptation mechanism can effectively release congestion before packet loss occurs. This has increased overall perceived voice quality when compared with existing adaptation method. This lightweight feedback-free early adaptation mechanism can be easily applied to other mobile devices to improve VoIP quality in current NGN VoIP applications.

Index Terms—VoIP, Adaptation, IMS, SIP, QoS, PQoS, Android, UMTS, AMR.

I. INTRODUCTION

International Telecommunication Union (ITU) defines the fundamental difference between NGN and today’s network as the switch from current circuit switched networks to all IP packet based systems. This implies that the Internet is the crucial player in the rollout of NGNs. The dependence on the Internet causes NGNs to face operational implications especially in users’ quality of experience. This happens because of the best effort nature of the Internet which makes the quality of VoIP communication very unreliable. The available bandwidth, packet losses and delay which are crucial for real time voice and video communication are not guaranteed in the Internet. This challenge has prompted an extensive research in the field of VoIP quality adaptation.

The majority of the previous research in the VoIP quality has concentrated in the management of QoS at the network level such as Differentiated Services (DiffServ) and Multi-Protocol Label Switching technologies (MPLS). But these technologies, despite their capabilities in reducing transmission delays and network congestion in the Internet, have failed to fully solve the problem of VoIP quality issues because the network equipment installed with QoS protocols are not widely available, and it is impractical and expensive to manage every call in real time over the Internet.

The difficulties in managing VoIP quality at network levels triggered the start of the end user initiated approaches, the most popular approach is the use of multiple speech coding techniques that provides a choice of a codec for a particular network bandwidth requirement (e.g., from 64 Kbps PCM to 13 Kbps GSM). Another approach is the use of adaptive codecs such as Adaptive Multi-Rate codec (AMR) which can adapt send bit rates from 4.75 Kbps to 12.2 Kbps according to network conditions.

The importance of VoIP quality adaptation stems from the fact that end users are the best judges of VoIP quality, therefore, satisfying users with quality of offered VoIP services by network and service providers has huge potential in retaining users, hence avoiding churn and increasing revenue.

There are many VoIP quality adaptation schemes in the literature which can be classified into two categories in general. The first one is to adapt send bit rate according to packet loss information [1]. The second one is to adapt send bit rate according to PQoS metrics (e.g. Mean Opinion Score (MOS)) derived from packet loss or delay [2] or combined packet loss and PQoS metrics [3]. The calculations of packet loss rate or MOS metrics are normally carried out at receiver side and relevant information are sent back via feedback channel (e.g. via RTCP report). These methods are based on packet loss information which is an explicit reflection of network congestion. The use of feedback mechanisms causes further delay in reaction. This results into an adaptation which is normally too late as loss and sometimes burst losses would have already happened and their impact on perceived voice quality is inevitable.

In this paper, we propose a novel Early VoIP Adaptation Scheme (EVQAS) which can adapt send bit rate (by switching AMR modes) based on the buildup of queuing delay. This proactive approach avoids congested packet loss and is also feedback-free, and it can easily be used in mobile/wireless VoIP applications. The paper also identifies the cause of packet losses in UMTS access network.

In order to present an NGN/IMS testbed that performs the proposed EVQAS, Open IMS Core (OIC) is used. Two
Android Dev 1 mobiles are used as IMS clients for VoIP calls registration, session initiation and termination over UMTS access network. The AMR codec is used, for which appropriate modes are selected based on the level of network congestion. The use of adapting AMR codec modes has its advantage over codec switching, as it is more bandwidth efficient because session parameters’ re-negotiation is not required.

Preliminary results have shown that EVQAS has advantage over the existing schemes as it releases network congestion before packet losses occur.

The rest of this paper is organized as follows. Section II outlines the most relevant related work. Section III presents the testbed and the experimental setup. Feedback-free EVQAS is proposed in Section IV. Experimental results and evaluation of EVQAS are described in Section V. Future work and conclusions follow in Section VI.

II. RELATED WORK

There exists a number of VoIP quality adaptation schemes in the literature, the most recent and relevant one has been proposed by Myakotnykh and Thompson in [4]. The scheme uses two parameters, instantaneous quality level computed by E-model [2] and perceptual metric which estimates the integral speech quality as inputs to adaptation decision. The scheme, though extensively analyzed in simulation, is based on E-model which is limited to packet loss information and delay.

Our previous work [5] reported an Open IMS Core testbed with VoIP quality adaptation using Android platform as IMS clients with two modes of AMR codec. Though the adaptation mechanism was implemented with the purpose of showing adaptation concept, extra traffic was introduced by the use of SIP UPDATE feedback mechanism in order to change AMR codec modes.

Our work in [6] proposed the first adaptive VoIP scheme that uses an interval Type-2 fuzzy logic which infers network state from average delivered perceived quality of service and its degradation due to network congestion and updates an AMR codec mode to match voice quality to available network bandwidth. Although the scheme achieved robust performance in the presence of input imprecision, early adaptation scheme was not considered.

Bilbao et al. [7] showed the relevance of implementing a dynamic and real time VoIP service adaptation mechanism in order to avoid PQoS degradations by investigating the most relevant parameters affecting the PQoS experienced by end users in VoIP services over UMTS. The detailed simulation study introduced the use of SIP re-Invite as a feedback mechanism during adaptation of AMR modes. SIP re-Invite messages have adverse effect because it introduces more delay and traffic during adaptation.

As shown above, all adaptation schemes’ decisions are only triggered when there is noticeable VoIP quality degradation. This implies that the adaptation actions, given the network condition are too late.

Network congestion is the major cause of VoIP quality degradation, which is revealed by queueing delay and packet loss. Ellis et al. [8] carried out experiments to characterize IPTV-lite traffic on residential links. The findings grouped packet loss in congestive and non-congestive losses, it was observed that congestive losses are the major cause of quality degradation and normally occur in bursts after queue overflow. The non-congestive losses occur due to link access errors which are random in nature and can be easily rectified by FEC-based algorithms. It was suggested that a mechanism that can detect the increase in delay due to queue buildup and reduce the bottleneck traffic rate are desirable to avoid queue overflows.

Ngamwongwattana and Sombun [9] used observations of the delay and packet loss correlation in order to characterize congestive packet loss. The proposed scheme distinguished between packet loss due to a limited bandwidth bottleneck and random packet loss due to high level statistical multiplexing. Though simulation results are promising, but the implementation at the end users is very limited in terms of finding queueing capacity and delay of a bottleneck link in one way delay.

This paper proposes a Feedback-free Early VoIP Quality Adaptation Scheme (EVQAS) for VoIP voice calls under IMS by using UMTS access network. EVQAS makes decision to adapt AMR codec modes before the congestive packet losses occur. EVQAS is based on RTP packet dispersion gaps method to detect the buildup of queueing delay. This method is simple to implement over the end to end VoIP devices, and does not need feedback mechanism. It eliminates the inherent inaccuracy of one way delay synchronization by GPS or NTP. The CBR characteristics of voice packets, constant packet sizes and packetization scheme, make RTP packet dispersion gaps method preferable in detecting congestive packet losses. It is non-intrusive and only measures RTP packets dispersion by themselves.

The packet dispersion technique to estimate link capacity was proposed by [10] and has recently been used in [11] for rate adaptive video streaming. The simulation results in [11] showed that by dynamically adapting the bitstream output of a transcoder or video encoder to a rate is less likely to lead to packet loss. The feedback mechanism and the variability of video packet sizes can be a major drawback in the implementation of dispersion technique used in [11]. The feedback mechanism can lead to late adaptation actions and by itself can cause quality degradation.

III. EXPERIMENTAL TESTBED

Figure 1 depicts the overall testbed developed to implement EVQAS in NGN. The testbed has Open IMS Core for RTP session establishment and termination. Two Android Dev 1 mobiles were used as IMS clients for UMTS access networks. Hutchison 3G UK Limited provides the UMTS access network.

The testbed is based on IMS architecture. IMS aims at the convergence of mobile, wireless and fixed broadband data networks into common network architecture where all types of data communication are hosted in all IP environments.
using session initiation protocol (SIP). Open IMS core is used because it is an open source implementation of IMS.

The Android Dev 1 mobile has been chosen because it is built with Android SDK. The Android platform is an open software stack for mobile devices including an operating system, middleware and key applications. Developers can develop their own applications in Android without seeking approval from Google. Android Dev 1 mobile is also capable of connecting to UMTS, WLAN, EDGE and Bluetooth access networks. In this testbed G1 mobile is used with UMTS access networks.

To be compatible with IMS architecture, Android is ported with SIP and RTP stacks. SIP stack is light weight MJSip for mobile applications and for that of RTP is based on JLibRTP. The SIP and RTP stack were modified to include monitoring and adaptation functions for VoIP quality adaptation.

![Android Dev 1 handset architecture with RTP and SIP stacks](image1)

**Fig. 2.** Android Dev 1 handset architecture with RTP and SIP stacks

The ported RTP and SIP stacks are depicted in Figure 2. The SIP Manager is the main entry for all SIP methods, it provides wrapping of the SIP stack functionalities and implements SIP listener interface that handles incoming and outgoing SIP messages. The SIP Manager interacts with the Media Manager for triggering RTP sessions. The Media Manager is responsible for receiving and transmitting RTP session through the AV Rec/Trans bloc.

The development under Android Dev 1 handset was carried out by using Android SDK 1.5 release 2 under the Eclipse IDE plugged in with Android Development Tools (ADT). At the moment Android Dev 1 handset only provides access to PCM audio format for RTP session, due to this limitation open source Opencore-amr was compiled and ported into the RTP stack.

### IV. FEEDBACK-FREE EARLY VOIP QUALITY ADAPTATION SCHEME

EVQAS depends on RTP packets dispersion gap technique to detect the buildup of queuing delay in order to adapt AMR modes before RTP packets are dropped due to queue overflow. RTP packets in voice calls are periodically transmitted with constant packet sizes, and by capturing the packets, the receiver can infer the network state. Within RTP packets headers, the receiver will be able to know the codec type, payload size and if it is AMR, the codec mode will be extracted from the payload header. From the AMR codec characteristics, the packetization delay which defines the inter-departure time is by default set at 20 ms. The transmission rate, for the case of AMR122 is 12.2 Kbps.

The dispersion gap is defined as,

$$\delta = \max(D_{pack}, D_{pack} + D_q, D_{trans} + D_{prop} + D_{proc})$$ \hspace{1cm} (1)

where $D_{pack}$, $D_q$, $D_{trans}$, $D_{prop}$ and $D_{proc}$ are delays caused by codec packetization, queues, transmission, propagation and processing, respectively. Processing delay may constitute routers and two end to end devices processing delays. With the assumption that the RTP packets will follow the same path during an ongoing session, $C = D_{trans} + D_{prop} + D_{proc}$ will be constant and can be ignored and hence (1) becomes,

$$\delta = \max(D_{pack}, D_{pack} + D_q)$$ \hspace{1cm} (2)

By knowing $D_{pack}$ from the payload type, $D_q$ can easily be computed because $\delta$ is calculated at the receiver side denoting the interval in which the arriving RTP packet is dispersed from its inter-departure gap (e.g., 20 ms for AMR).

Ngamwongwattana and Thompson [12] came up with the same equation but $E$ was introduced as an epoch of excursion of queueing delay for time synchronization. From the empirical results conducted in this paper, it is impractical to obtain $E$ unless the environment is in a steady state, this contradicts with VoIP calls which are usually of a short time in nature.

In order to detect gradual buildup of queueing delay and hence network congestion, EVQAS uses (2). EVQAS is initially trained by real time RTP packets in order to identify the maximum value of $D_q$ at which RTP packets are beginning to be dropped (congestive packet losses) by the queue. EVQAS ignores the non-congestive packet losses because they can easily be rectified by FEC-base algorithms, non-congestive losses will be known by observing packet losses without a gradual buildup of $D_q$. 

---

978-1-4244-5637-6/10/$26.00 ©2010 IEEE
Square coefficient of variation (SCV) is used to calculate the level of congestion as it is a dimensionless measure of dispersion. SCV is defined as,

$$SCV = \left( \frac{\sigma}{\mu} \right)^2$$

(3)

where the mean $\mu > 0$ and $\sigma$ is the standard deviation of RTP packet dispersion gaps. SCV has been commonly applied in queueing theory for queue length distributions. In this paper, it has been observed that $0 < SCV < 1$, with $SCV \to 1$ as being heavily congested network and $SCV \to 0$ as being lightly congested network. The empirical results have shown that the gradual buildup of queueing delay with the corresponding $SCV > 0.37$ resulted into congestive packet losses. In this regard, a threshold of SCV was set at 0.3 for EVQAS to perform adaptation actions. Further research on SCV threshold is open for different network types and environment.

This paper uses the MOS model proposed by Lingfen and Ifeachor [3] for end to end conversational QoE measure for AMR codec. AMR122, AMR795 and AMR475 modes are used, the choice is justifiable by the range of send bit rates that would be expected to make noticeable change in releasing network congestion during adaptation procedure.

The feedback-free concept is due to the fact that the receiver adapts its sending AMR mode after detecting the gradual buildup of queueing delay, this mode will be picked up by the sender through the payload header and in return the sender adapts its sending AMR mode in order to sync with receiving mode. Therefore, this process does not require conventional feedback mechanisms such as RTCP. This process is in line with AMR 20 ms inter-departure time where as the default RTCP report is sent at 5000 ms periodically.

V. EXPERIMENTAL RESULTS AND EVALUATION

Figure 3 depicts one of the queueing delay traces collected in one of the Android Dev 1 phones during a VoIP call with AMR122 mode. Queueing delay traces were collected for a week during working hours, and occurrences of the queueing delay buildup can easily be seen in the trace. EVQAS has not been employed on this trace.

This section also distinguishes two observed packet loss events, i.e., congestive and non-congestive packet losses. It can be seen in Figure 4 (a cross section of Figure 3) that congestive packet losses occurs when there is a gradual buildup of the queueing delay, this is then followed up by the sudden fall of the queueing delay back to its normal delay. The packet losses occur when the queue is full and start to overflow. The non-congestive packet losses occur randomly and independent of the queueing delay buildup. This is illustrated in Figure 5 (a cross section of Figure 3). This can be caused by the link layer errors of the UMTS access network.

With the introduction of the proposed EVQAS and the use of SCV as the measure of congestion level, the buildup of queueing delay is bounded at approximately 0.6 ms in order to avoid congestive packet losses due to queue overflow, this is depicted in Figures 6 and 7, both showing traces of queueing delay with EVQAS in place.

The cumulative distribution functions (CDF) of queueing delays are illustrated in Figure 7. It is clearly shown that with EVQAS, only 0.5% of queueing delay occurrences where above 0.4 ms with the maximum of 0.6 ms in which there was no congestive packet losses because the queue did not overflow. This is preferably compared to the one without EVQAS where by 2% of queueing delay occurrences were above 0.4 ms with the maximum of 0.93 ms at which congestive packet losses occur.
losses were observed due to queue overflow. The two CDFs have clearly shown the overall improvement of queueing delay when EVQAS is in place. Both CDFs curve fittings resulted into \( f(x) \) with EVQAS favorably avoiding up to 60% of queueing delays and hence congestive packet losses.

EVQAS is also compared to the adaptation scheme of [5], the corresponding MOS values are depicted in Figure 8, initially AMR122 was in use in both schemes, and after detecting gradual buildup of the queueing delay with the use of SCV, EVQAS adapted AMR122 to AMR795 in order to release congestion. After several seconds, the queueing delay dropped down to its base line value, this was also detected by SCV which enabled EVQAS to adapt the VoIP dialogue with a better quality AMR122. It is clearly seen that the quality in [5] rapidly deteriorates before adaptation has occurred.

VI. CONCLUSIONS

This paper has proposed a novel feedback-free early VoIP quality adaptation scheme by using RTP packet dispersion gaps in order to detect congestive packet losses. The congestion level was determined by the squared coefficient of variation. The proposed scheme was implemented, tested and evaluated under IMS with UMTS access networks connecting two Android Dev 1 phones. The preliminary results have shown that EVQAS has an advantage over other schemes in the literature in which adaptation actions are too late and always cause noticeable deterioration in quality to end users. EVQAS uses AMR codec modes whose adaptation actions do not require IMS dialogue parameters’ re-negotiation or feedback mechanism and therefore meet 3GPP requirements of seamless service continuity.

Future work will include extensive analysis and comparisons with other adaptive codecs such as SILK and SPEEX. A detailed analysis and evaluation of SCV will also be carried out under different network environments.

REFERENCES