Computer Networks 52 (2008) 2473–2488

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

Adaptive playout scheduling algorithm tailored for real-time packet-based voice conversations over wireless ad-hoc networks

Sofiene Jelassi *, Habib Youssef

Research Unit PRINCE, Higher Institute of Computer Science and Technologies of Communication, Hammam Sousse, Tunisia

ARTICLE INFO

Available online 22 April 2008

Keywords:
Multimedia over multi-hop wireless ad-hoc networks
Playout algorithms
User satisfaction
E-Model

ABSTRACT

The effective provision of real-time, packet-based voice conversations over multi-hop wireless ad-hoc networks faces several stringent constraints not found in conventional packet-based networks. Indeed, MANETs (mobile ad-hoc networks) are characterized by mobility of all nodes, bandwidth-limited channel, unreliable wireless transmission medium, etc. This environment will surely induce a high delay variation and packet loss rate impairing dramatically the user experienced quality of conversational services such as VoIP. Indeed, such services require the reception of each media unit before its deadline to guarantee a synchronous playback process. This requirement is typically achieved by artificially delaying received packets inside a de-jitter buffer. To enhance the perceptual quality the buffering delay should be adjusted dynamically throughout the vocal conversation.

In this work, we describe the design of a playout algorithm tailored for real-time, packet-based voice conversations delivered over multi-hop wireless ad-hoc networks. The designed playout algorithm, which is denoted MAPA (mobility aware playout algorithm), adjusts the playout delay according to node mobility, which characterizes mobile ad-hoc networks, and talk-spurt, which is an intrinsic feature of voice signals. The detection of mobility is done in service passively at the receiver using several metrics gathered at the application layer. The perceptual quality is estimated using an augmented assessment approach relying on the ITU-T E-Model paradigm while including the time varying impairments observed by users throughout a packet-based voice conversation. Simulation results show that the tailored playout algorithm significantly outperforms conventional playout algorithms, specifically over a MANET with a high degree of mobility.

1. Introduction

A mobile ad-hoc network is a collection of heterogeneous nodes, which move randomly within an area and communicate with each other via wireless interfaces. Contrary to conventional last-hop wireless networks, MANETs enable the establishment of multi-hop wireless connections between any source–destination pair without the use of any existing infrastructure. In fact, inside a MANET all nodes act as routers and forward received packets to nodes within radio range.

MANETs are challenging networks since they exhibit several characteristics not found in conventional packet-based networks. For instance, the shared nature of the wireless channel entails an end-to-end effective data rate of few kilobits, despite the quick rise of point-to-point data rate which reaches currently 54 Mbps for some wireless cards and norms such as IEEE 802.11g and IEEE 802.11a [1]. This is mainly due to the interference problem between Intra and Inter flows in the same vicinity while crossing multiple wireless hops. In fact, Chen et al. proved that the maximum spatial reuse is roughly equal to 1/4 of

* Corresponding author.
E-mail addresses: Sofiene.Jelassi@infcom.rnu.tn (S. Jelassi), Habib.Youssef@fsm.rnu.tn (H. Youssef).

1389-1286/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2008.04.009

chain length for particular chain topology and node parameters [2]. This reduction of end-to-end data rate is also due in part to the control message handshake done by CSMA/CA to avoid hidden and exposed terminal problems [2]. Moreover, the problems related to the transmitted signal such as fading, or reflection, cause a high packet and bit error rates. More critically, the mobility of all nodes entails a high link change rate which emphasizes the instability and blackout of used paths during a session.

The MANET properties make the provision of packet-based conversational applications such as packet based voice conversation a challenge. Indeed, such applications require bounded end-to-end delay and delay jitter, but can allow a limited packet loss ratio. Packet losses occurring during a voice conversation impair intelligibility while delay impairments affect interactivity. Despite the challenging constraints imposed by MANETs, adaptive behavior of multimedia applications at sender and receiver sides can improve dramatically the perceptual quality. For instance, a recently proposed CODEC reported in [3] achieved a packet loss rate tolerance of as much as 30%. Moreover, with the intelligent management of a buffer with appropriate size at receiver side, delay jitters of as much as 50 ms may be tolerated. This buffer is commonly called de-jitter buffer.

A feature of delay-sensitive conversational applications is the requirement to receive each media unit before its computed deadline. The breaking of this constraint will result in the occurrence of gaps, significantly impairing the perceptual quality. To reduce gap occurrences, the receiver uses a de-jitter buffer where packets are temporarily stored before sending them to the decoder. This extra delay is introduced artificially to compensate the variable network delays. However, finding an optimal equalization value for real-time applications is unfeasible since the network behavior is unpredictable. For real-time conversational applications, a large equalization delay is preferred (e.g., 2–3 s) to reduce the occurrence of late arrivals. Unfortunately, such large end-to-end delay will deteriorate the perceptual quality by affecting the interactivity between communicating parties. Consequently, the equalization delay, referred to commonly as the playout delay, should be adapted dynamically throughout a packet-based voice conversation to minimize both the end-to-end delay and packet losses due to late arrivals. This trade-off is achieved using a playout algorithm computing the deadline of each received packet. Playout algorithms of voice packets over the Internet have been extensively studied in the literature [4,5]. Playout algorithms can be either static or adaptive. Static playout algorithms retain a constant playout delay during a vocal conversation. However, recommended adaptive playout algorithms dynamically adjust the playout delay throughout a vocal conversation. Existing adaptive approaches fall into two categories:

– *Network-based approach*: It adapts the playout delay by monitoring a set of performance objective metrics measured at application layer such as the network delay, the de-jitter buffer depth, or the packet loss rate. The adaptation instant and its level are independent of the played signal structure.
– *Application-based approach*: It adapts the playout delay according to the measured performance metrics, but considers the features of the played signal such as talk-spurts in the case of voice conversations and type of frame in the case of video conversation.

In this paper, we describe the design of a tailored voice packet playout algorithm, denoted MAPA, to suit network dynamics incurred by voice packets over a MANET. To this end, a comprehensive analysis of the key parameters of a voice conversation, namely one-way network delay and packet losses, is performed in order to adequately characterize the effect of mobility at application layer. The designed algorithm MAPA adjusts dynamically the playout latency of received voice packets according to mobility-induced path switching, which is the main feature of MANETs, and talk-spurt, which is an intrinsic characteristic of voice signals. During path switching periods, voice packets are played cleverly on a per-packet basis while maximizing the perceptual quality as much as possible. However, during normal periods, MAPA plays voice packets according to a baseline per-talk-spurt playout algorithm which is extended to reduce the distortion effects due to severe compression and expansion of original silence period duration. The mobility is detected passively at the application layer by the receiver using the inter-packet delay difference coupled with packet out-of-order ratio metrics. Mobility-induced path switching is declared once measured parameters exceed a set of empirically calibrated thresholds. The rating factor is used to evaluate the performance of MAPA. This metric is estimated objectively using an extended computational assessment algorithm based on the E-model approach.

The remainder of this paper is organised as follows. Section 2 gives an overview of conventional voice packet playout algorithms designed for voice conversations running over a wide area IP network. Section 3 provides an analysis of key parameters of voice conversations over a MANET. Section 4 describes the algorithm MAPA, designed specifically to play voice packets transferred over a MANET. Section 5 describes the approach integrated in MAPA to detect mobility at the application level. An extended methodology to assess the perceptual performance of MAPA over a connection with time varying impairment is given in Section 6. Performance results and discussion are given in Section 7. Finally, we conclude in Section 8.

2. Related work

Real-time packet-based voice conversations require bounded end-to-end delay and delay jitter. Fig. 1 shows the temporal constraints linking the sender, transfer, and playout processes of a typical packet-based voice conversation transferred over a packet-based, best-effort network. The form of the staircase is due to the transmission of one packet every 20 ms, each of length 160 bytes. The generated traffic imitates the ITU-T G.711 codec output while disabling the voice activity detector algorithm, abbreviated as VAD. As shown in Fig. 1, voice packets reach the receiver side with variable network delays. The figure