Connectivity Aware Instrumental Approach for Measuring Vocal Transmission Quality over a Wireless Ad hoc Network

Sofiene JELASSI and Habib YOUSSEF
Higher Institute of Computer Science and Communication Technologies
Research Unit PRINCE
Hammam Sousse, Tunisia
E-mail: sofiene_jelassi@yahoo.fr and habib.youssef@rnu.fsm.tn

Abstract-**This paper describes an on-line, end-to-end and non-intrusive approach to estimate the vocal transmission quality over wireless ad-hoc networks. The proposed assessment algorithm, denoted PEVOM (Perceptual Evaluation of Voice over MANETs), is based on the fact that the observed quality of packet based voice conversations is time varying and that mobile users may tolerate degradation of perceptual quality and even the interruption of service. The evaluation of transmission quality adopted by PEVOM is relying on E-Model paradigm. PEVOM includes an original extension to E-Model to adequately rate the distortion effect due to the intermittent nature of established connection incurred by mobile users over a MANET. Specifically, PEVOM checks passively the connectivity between communicating parties by monitoring periodically the ratio of active to total time. When mobile users are connected, PEVOM calculates the perceptual quality using the conventional E-Model. However, when mobile nodes are disconnected, PEVOM calculates the perceptual quality according to an expected recency effect. At the end of each connected/disconnected period, PEVOM estimates the perceptual quality of each interval using a weighted average over time. At the end of a vocal session, PEVOM rates the overall conversational quality using a weighted average over time of the produced rating factor of each connected/disconnected interval. The rating behavior of PEVOM is verified through a study of the vocal transmission quality over an intermittent multi-hop wireless connection.**

I. INTRODUCTION

A mobile ad-hoc network (MANET) is a set of mobile nodes communicating with each other in peer-to-peer manner. In contrast to infrastructure based networks, transfer in a MANET is done without relying on any pre-existing infrastructure. Indeed, in a MANET, a terminal node should assure traffic forwarding for other nodes. This results in the establishment of multi-hop wireless connections. MANETs are very attractive for many applications and environments such as conference and convention center communications as well as emergency situations such as law enforcement and military activities.

It is highly desirable to support interactive voice conversation over a MANET to satisfy users' requirements in several scenarios. For this reason, many solutions are suggested in the literature in order to improve the quality of experiment (QoE) incurred by mobile users [1, 2, 3]. Thus, assessment of QoE is used in order to evaluate the performance of the proposed solutions. The accuracy of the assessment procedure estimating the QoE is important since the obtained results are used to classify the suggested solutions and adopted technologies. Indeed, the assessment procedure has a fundamental economical aspect specifically over telephony networks where it is used to sort service providers. Hence, the used methods of assessment should judge fairly the users opinion. Due to the diversity of impairment sources and psychological aspect of receivers, the assessment procedure can be very complex and expensive specifically over wireless networks.

The assessment of a voice conversation at user level can be done subjectively or objectively. Subjective tests derive the rating factor according to the opinion score of a set of human subjects using a set of emulated scenarios. Specifically, these subjects are placed in a lab environment to evaluate a set of voice files (Listening Quality) or conversational experience (Conversational Quality). The evaluation process is done using a dedicated scale. For instance, the standard metric used to evaluate the heard speech file is the Mean Opinion Score (MOS). The MOS values vary between 1 (bad quality) and 5 (excellent quality). According to the collected MOS values of an experiment, the final MOS of the analyzed sequence is derived based on statistical analysis. The ITU-T recommendation number P.800 describes precisely how these tests should be conducted [4]. The major drawback of subjective methods is the requirement to re-iterate these tests for each new codec, network strategy, playout algorithm, or recovery method in order to quantify the perceived quality. These tests are unable to rate the perceptual quality in real time in order to monitor live conversations. Furthermore, subjective approach is time consuming, cumbersome, and expensive particularly to assess mobile wireless networks.

To circumvent subjective approach drawbacks, objective approach is used as alternative. Indeed, objective assessment techniques rate the perceptual quality using algorithms and formulas implemented either on terminal or intermediate nodes. There are several objective methods assessing live voice traffic in real time [5, 6, 7]. These properties are attractive for several applications and scenarios such as prediction, monitoring, and diagnosis. We should note here that an objective approach should be calibrated using subjective results in order to produce a reliable feedback. In recent years, several objective assessment methods for measuring MOS have been proposed and developed [8, 9].

QShine 2008, July 28-31, 2008, Hong Kong, Hong Kong.
 ISBN 978-963-9799-26-4
DOI 10.4108/ICST.QSHINE2008.4090

Particularly, the ITU-T recommendation number G.107 defines E-Model, a computational model-based parametric assessment algorithm combining all impairment sources that affect a voice conversation to produce a single rating factor, which can be easily converted into a MOS scale [8].

This paper describes an original assessment algorithm to rate the transmission quality of voice conversations on call-by-call basis over a MANET. The proposed algorithm is denoted PEVOM, as abbreviation of Perceptual Evaluation of Voice over a MANET. PEVOM relies on ITU-T E-Model paradigms. PEVOM models adequately the loss of connectivity events to estimate the overall transmission quality. In fact, users in multi-hop ad-hoc networks can tolerate the degradation and even the interruption of service due to the allowed high mobility. The parametric and non-intrusive nature of PEVOM enables it to monitor and manage live voice sessions carried over a MANET.

The remainder of this paper is organized as follows: Section 2 gives a quick overview of E-Model. In Section 3, we present an ETSI extension to E-Model to rate voice conversation transmitted over a time varying impairment networks. In Section 4 we describe the main building blocs of PEVOM designed specifically to rate live packet based vocal communications transmitted over multi-hop wireless connections. In Section 5, we discuss some issues related to the implementation of PEVOM and we give an illustrative scenario where PEVOM can be exploited. We conclude in Section 6.

II. THE E-MODEL

E-Model is a parametric assessment approach proposed initially for planning purposes to predict the conversational quality of vocal services over telephone networks [8]. E-Model rates the transmission quality by combining sources of impairment experienced on the mouth to ear (M2E) path, and providing as output a rating factor, denoted R. The rating factor is a scalar ranging from 0 to 100 corresponding respectively to the worst and the best transmission quality. A rating factor value below 60 corresponds to an unsatisfied transmission quality [8]. E-Model does not require the original voice sequence or the degraded voice sequence like intrusive approaches to rate the transmission quality, but needs the knowledge of specific measures computed over the M2E path [10]. These measures are provided as input to the assessment algorithm in order to compute the transmission rating factor. E-Model is currently under extensive study in order to enhance its performance over several kinds of transport networks [10].

E-Model is based on the assumption that impairment factors are additive on psychological scale. This simplification, which is true to a certain extent, is needed to make the model tractable. According to the ITU-T recommendation number G.107, the rating factor R can be obtained through Equation (1).

$$R = R_0 - I_s - I_e - I_d + A \qquad (1)$$

where, R_0 represents the optimal transmission quality stemming from the intrinsic distortions due to the basic signal-to-noise ratio including the effects of noise sources such as room and circuit noises, I_s captures the effects of impairments occurring simultaneously with the speech signal such as quantification and attenuation, I_e captures the distortions related to the equipments used during the interaction such as low bit rate codec and introduced losses occurring over VoIP systems, I_d models the impairments related to the propagation delay, and A represents an advantage factor that accounts for user willingness to accept some quality degradation in return for ease of access (e.g. cell phone). The value of A varies between 0 and 20 corresponding respectively to a wired network and 2 satellite hops [8].

Actually, the calculation of the R factor is based on 21 input parameters and includes complex mathematical formulas. To simplify the calculation, the ITU-T has recommended 14 default parameters that are independent of the transmission quality of the network used to carry vocal conversations. These parameters represent the basic signal-to-noise ratio, the simultaneous impairments I_s and the advantage factor A. In the context of real-time packet based voice conversations over a Best Effort (BE) network, the delay impairment, I_d, and equipment impairment I_e, are pertinent. By selecting the recommended default values (R_0 = 94.77, I_s = 1.41), the initial equation to derive R is reduced to the following Equation [10]:

$$R = 93.2 - I_d(T_a) - I_e(\text{codec}, \text{loss}) + A \qquad (2)$$

where I_d has as input the average one-way absolute propagation delay and I_e has as input the codec used and the packet loss rate occurring during a vocal session. Now, it is required to model the distortion effects due to delay, I_d, and equipment, I_e, to compute the rating factor. Typically, these models are derived and calibrated using subjective experiments. The distortion effects of I_d are well known and easily modeled [5]. This factor includes the distortion due to talker and listener echoes as well as the absolute delay T_a. By assuming a perfect echo canceller, the distortion due to propagation delay can be estimated using Equation (3). This equation is derived using a linear regression model applied on subjective results [7].

$$I_d(T_a) = 0.024\, T_a + 0.11\, (T_a - 177.3)\, H(T_a - 177.3) \qquad (3)$$

$$\text{where} \begin{cases} H(x) = 1 \text{ if } x < 0 \\ H(x) = 0 \text{ if } x \geq 0 \end{cases}$$

The absolute propagation delay T_a is the sum of the transit time inside the network and processing time at each terminal. This delay includes among others framing, buffering, and CODEC delays.

In contrast to I_d, the distortion effects captured by I_e depend mainly on the codec used and loss distribution which can be either random or bursty [5]. Naturally, a model for each configuration should be derived subjectively. These experiments should be re-iterated for each codec and loss pattern. The wide range of academic and commercial voice codecs and loss distribution make these experiments unpractical, expensive, and time consuming. To avoid subjective experiments, S. Lingfen et al have proposed in [11] an objective method enabling to estimate I_e for any codec and loss pattern. In order to derive the adequate model of I_e for each codec with respect to packet loss rate, a regression process should be applied. This model can be used on-line to estimate the degradation due to equipments. Basically, developed models have the following form: $I_e = a + b \times \ln(1 + c \times loss)$ where a, b, and c are real coefficients obtained via a logarithmic regression. For instance, the adequate coefficients of G.711 codec with packet loss concealment are a = 0, b = 30, and c = 15 [7]. A similar approach has been recently standardized in the ITU-T recommendation number P.834.

III. THE ETSI EXTENDED E-MODEL

To use the E-Model as an on-line tool for quality assessment of live packet based voice conversations, several adaptations should be made. Indeed, E-Model is intended to predict the transmission quality of vocal services carried over circuit switched networks which are characterized by a constant propagation delay and a negligible loss rate distributed uniformly during a vocal conversation. However, real-time packet based voice communications over IP networks are characterized by the varying nature of impairments over time. Indeed, real-time traffic over IP network suffers from important delay jitters and varying loss patterns. In fact, it is well verified empirically that packet losses over an IP network occur in bursts [12].

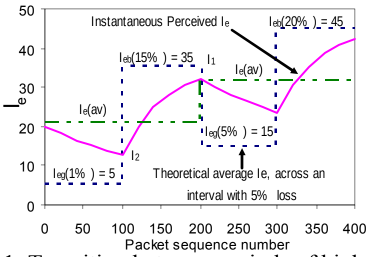


Figure 1: Transition between periods of high and low loss. Theoretical vs. Instantaneously perceived I_e

To incorporate the perceptual effects of the time varying nature of impairments, the ETSI extension of E-Model introduces new concepts such as "instantaneous quality" to denote the measured or calculated quality, "perceived quality" to denote the perceptual quality that would be reported by the user at some time during the vocal conversation, "time varying packet loss behavior" to denote the oscillation between "gap" and "burst" loss periods characterized respectively by a low and high packet loss rate, and "recency effects" which incorporate the distortion effect due to a specific distribution of impairments [5]. A well-recognized recency effect can be described as follow: when the instantaneous theoretical quality changes from "good" to "bad" at some moment during the call, then it can be expected that initially users would not be too concerned about this. However, after some time, the listener becomes annoyed with the voice quality degradation. The same phenomenon is expected and observed when the instantaneous theoretical quality switches from "bad" to "good" [13]. This effect is modeled using exponential curves with a time constant of 5 sec in the transition from "good" to "bad" and 15 sec in the transition from "bad" to "good" [5]. Figure 1 illustrates the transitions between periods of "high" and "low" loss which corresponds respectively to a "bad" and "good" periods.

A well-known extension proposed to assess the time varying packet loss behavior captured by I_e is described in the Annex E of an ETSI recommendation and integrated within the distribution of the commercial tool VQmon (Voice Quality Monitoring) [5, 13]. To capture the alternating behavior of packet loss during a vocal conversation and to provide on-line assessment facility, ETSI recommends using a 4-state Markov chain. The four states represent the condition of receiving or losing a packet within a high packet loss (burst) or low packet loss (gap) conditions. A packet loss event driven approach was used to compute the transition probabilities between connected states. At the end of the assessment period, a set of relevant information can be extracted from the model such as the average gap (g) and burst (b) durations, the density of packet losses under gap and burst conditions, the successive lost packet distribution, etc. These metrics are used to compute I_e factor separately for burst (I_{eb}) and gap (I_{eg}) periods. I_{eb} and I_{eg} correspond to the instantaneous I_e factors over the monitored interval which can be obtained using subjective curves (see Figure 1). The average perceptual quality $I_e(av)$ is given by:

$$I_e(av) = \left(bI_{eb} + gI_{eg} - t_1(I_{eb} - I_2)\left(1 - e^{-b/t_1}\right) + t_2(I_1 - I_{eg})\left(1 - e^{-g/t_2}\right)\right) / (b+g) \qquad (4)$$

where I_1 is the quality level at the change from burst condition to gap condition, I_2 is the quality level at the change from gap condition to burst condition, b and g correspond respectively to the average burst and gap durations expressed in seconds, t_1 and t_2 are two time constants set respectively to 5 and 15 sec (see Figure 1). The values of I_1 and I_2 are derived using Equations (5) and (6).

$$I_1 = I_{eb} - (I_{eb} - I_2)e^{-b/t1} \qquad (5)$$

$$I_2 = I_{eg} + (I_1 - I_{eg})e^{-g/t2} \qquad (6)$$

$I_e(av)$ value corresponds to I_e(codec, loss) in Equation (2). We note here that we can have several $I_e(av)$ values during a vocal call. For this reason, a weighted value of $I_e(av)$ may be computed at the end of a voice conversation where weights correspond to the duration of each observed $I_e(av)$ value. Moreover, it is empirically verified that distortions occurring at the end of a voice call reduce considerably the overall rating factor than distortions occurring at the beginning or the middle of a conversation. This reduction in perceptual quality is included while estimating the final equipment impairments, I_e(end of call) using a calibrated exponential curve.

$$I_e(\text{end of call}) = I_e(w) + (k \times (I_1 - I_e(av))) \times e^{-y/t_3} \qquad (7)$$

where $I_e(w)$ corresponds to the weighted value of $I_e(av)$, y represents the time delay since the last burst, t_3 is a time constant of typically 30 to 60 seconds and k is a constant set to a nominal value of 0.7. The user R factor can be computed as follows:

$$R = 93.2 - I_e(\text{end of call}) - I_d \qquad (8)$$

where I_d corresponds to the weighted average of each instantaneous I_d. The weights are the duration of each instantaneous I_d value observed during the call.

As we can see, the important reported extension does not include the effect of transient disconnections between the communicating parties. Indeed, this event is seldom observed over wired and last hop wireless networks. However, when voice conversations are carried over a MANET, disconnections become the rule. Therefore, the objective rating process should be changed to account for the occurrence of repeated disconnections. This constitutes the main contribution of this paper.

IV. PERCEPTUAL EVALUATION OF VOICE OVER MANETS (PEVOM)

In this section, we describe precisely how PEVOM rates the perceptual quality of conversations over a MANET. PEVOM follows the parametric assessment methodology. Therefore, the vocal perceptual quality is derived using models and specific measured parameters. Figure 2 depicts how PEVOM processes gathered measurements to estimate the transmission quality. Figure 2 illustrates the alternation between connected and disconnected states during a vocal conversation. We note that in this work, the word interval refers to a connected/disconnected period which can have a variable duration. PEVOM assesses independently the perceptual quality of each interval. This will produce a rating factor denoted R_i where i corresponds to the identifier of the i^{th} interval. PEVOM estimates the overall perceptual quality at the end of a vocal conversation using a weighted average of R_i values gathered during the conversation. The weights are the durations of the observed intervals. This methodology is recommended to objectively rate the perceptual quality over a time varying quality of a vocal sequence [14]. Formally, the rating factor at the end of a vocal conversation is estimated as follows:

$$R_{\text{end of call}} = \frac{\sum_{i=1}^{P} d_i \times R_i}{\sum_{i=1}^{P} d_i} \qquad (9)$$

where d_i corresponds to the duration of the i^{th} interval, R_i represents the estimated rating factor observed during the i^{th} interval, and P corresponds to the number of assessed intervals.

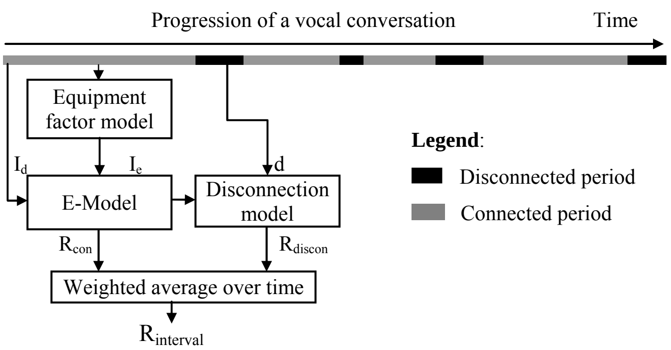


Figure 2: Methodology used by PEVOM to rate vocal transmission quality over a MANET

In order to compute the transmission quality of each interval, we rate separately the connected and disconnected periods. This will produce two measures of perceptual quality during connected and disconnected periods denoted respectively R_{con} and R_{discon} (see Figure 2). To derive the ranting factor of an interval, R_{con} and R_{discon} are weighted over time.

In order to derive R_{con}, we use the basic logarithmic models proposed in [7] and extended in [11] to produce the distortion due to equipment, denoted I_e. We also estimate the distortion due to delay according to Equation (3). The absolute delay used by I_d corresponds to the mean playout delay of each played packets. These sources of impairment are combined using Equation (2) to derive a value for R_{con} which estimates the perceptual quality during connected state.

On the other hand, R_{discon} is estimated using a recency effect. This requires the appropriate modeling of the effects of temporary service interruption on mobile users. During disconnected periods, we expect to observe the following recency effect: upon the occurrence of a disconnection we expect that mobile users will initially accept this impairment event since they are aware about the challenging properties of multi-hop wireless ad-hoc networks. However, beyond a certain threshold, users will be very annoyed and with high probability, the conversation will be abruptly broken. We propose to model this recency effect using an exponential decay. To this end, two parameters should be defined: the start point and the rapidity of decay. We choose as start point the rating factor estimated at the end of the previous connected period (R_{con}), and we express the rapidity of curve decay using a time constant. Formally, the exponential curve used to model the disconnection effect on transmission quality as a function of time is derived as follows:

$$R(t) = R_0 e^{-\frac{t}{\tau}} \qquad (10)$$

where τ represents the rapidity of decay expressed in unit of time, and R_0 corresponds to the last estimated rating factor of the previous connected period. To estimate R_{discon} at the end of a disconnected period, we compute the average rating factor by integrating over time R(t) which evaluates to the following expression:

$$R_{discon} = \frac{\tau}{d} \cdot \left(1 - e^{-\frac{d}{\tau}}\right) \cdot R_0 \qquad (11)$$

where d represents the disconnection duration. Certainly, the value of τ should be calibrated using subjective tests. In this work, we use the exponential curve plotted in Figure 3 which gives the rating decay factor to use for value of d varying between 0 and 8 sec. The time constant τ is set to 8.5 sec so that a disconnection of duration 6 sec results in a reduction of the rating factor to half of its initial value. We note here that the voice conversation is declared broken when the disconnection duration exceeds 12sec. According to our point of view, interested users will try to establish another voice session later.

Actually, it is required to detect the connectivity between communicating nodes to adequately rate the transmission quality. The connectivity detection can be either done in active or passive manner. An active connectivity detector, defined in the RFC 2786, sends one or several messages soliciting an urgent reply from the other side. When the timeout of all sent requests are triggered without any reply from the receiving node then the occurrence of loss of connectivity is declared.

The active strategy requires assessing the transmission quality on-line since the connectivity state is updated by a disconnection detector during the service using a background thread. Moreover, the requirement to inject live traffic will surely result in an overhead increase inside the network and at terminal nodes.

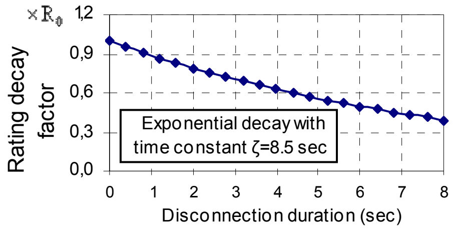


Figure 3: Modelling of disconnection impairment included into PEVOM

To circumvent active connectivity detection drawbacks, a passive disconnection detector may be used as an alternative. Passive methods base their decision on a recorded trace without any additional overhead. Basically, a passive disconnection detector computes a ratio, denoted CON, of active time to total time over a fixed window size. This monitoring window is shifted periodically by a specific duration. Then, according to a meticulously selected threshold, CON_0, the network state can be inferred. The network is assumed disconnected whenever the computed CON falls below CON_0. The threshold CON_0, the monitoring window size, and the shift duration should be carefully selected with respect to the specificities of a packet based voice stream. The receiver should be able to distinguish between silence periods and disconnected periods. Silence periods correspond to the temporary suspension of transmission at the source while disconnection periods correspond to path loss. Recall that silence suppression is performed at the sender side to reduce bandwidth consumption and to assist the playout algorithm. Note that passive approaches are relatively less accurate than active approaches, but in the context of this work the performance of passive strategy is sufficient. The current version of PEVOM implements only the passive approach.

Finally, we note that PEVOM is able to estimate the instantaneous quality. When mobile users are connected, PEVOM predicts the instantaneous transmission quality using Equation (1) applied on a window of a fixed duration. However, when mobile users are disconnected, PEVOM predicts the instantaneous transmission quality using Equation (10). This value can be reported on-line as feedback to mobile users for diagnosis purposes. We suggest an assessing period duration lying between 5sec and 10sec to accurately capture the time varying nature of impairments. Next, in the experimental section, we define the value of each used parameter.

V. PEVOM IMPLEMENTATION ISSUES AND SCENARIO UNDER MEASUREMENT

We discuss in this section some issues related to the implementation of PEVOM. Figure 4 illustrates the architecture of PEVOM. Basically, PEVOM probes voice packets reaching the receiver side in order to gather the required measures for transmission quality estimation such as packet losses and delay. As depicted in Figure 4, PEVOM considers the de-jittering buffer mechanism by including the buffering delay and losses due to late arrivals. These statistics can be provided directly to users for diagnosis and maintenance operations.

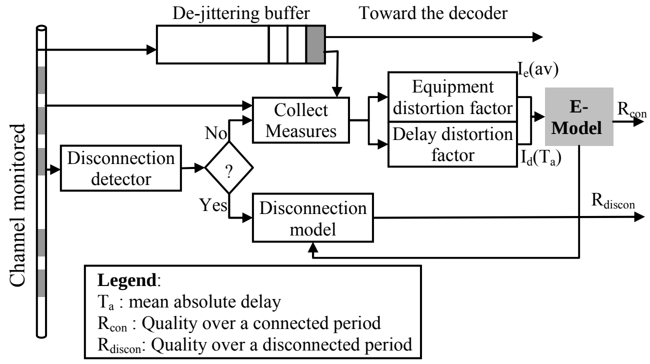


Figure 4 : Architecture of PEVOM

The disconnection detector works as a background process. The window size to detect disconnection is set to 3sec, the window is shifted by one packet, and CON_0 is set to 0.95. When PEVOM detects the occurrence of a new interval, we estimate the perceptual quality as described previously. At the end of the conversation, PEVOM returns the transmission quality over the vocal conversation.

This section aims to examine the rating behavior of PEVOM and to compare it with VQmon over a MANET. To this end, we simulate the multi-hop wireless network scenario depicted in Figure 5 using ns2-31 [15]. In the analyzed scenario, 6 mobile nodes are created and configured as follows. The transmission range of each node is selected so that each node can communicate only with its immediate physical neighbours. The five nodes A, B, C, D, and E are retained stationary during the simulation and only node F travels with a fixed speed according to the zig-zag trajectory shown on the figure from the point (120,70) toward the point (245, 160) and stops there. The trajectory makes node F intermittently loose communication with node E. A voice session is established between nodes E and F. This session is composed of two vocal streams following an ON/OFF model, one in each direction. During an active period, the source sends packets having a payload size equals to 160 bytes every 20 ms over UDP. This results in a data rate equal to 64 Kbps during active periods. The session duration is set to 120sec. We equip each receiver end with a de-jitter buffer managed according to the Algorithm 1 reported in [16].

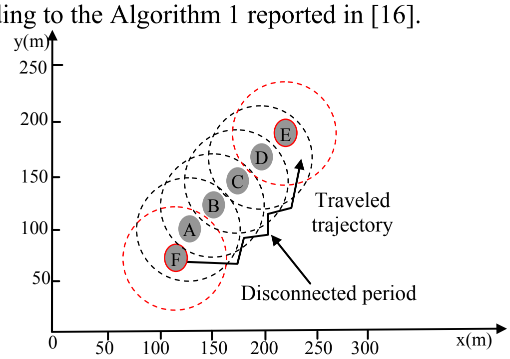
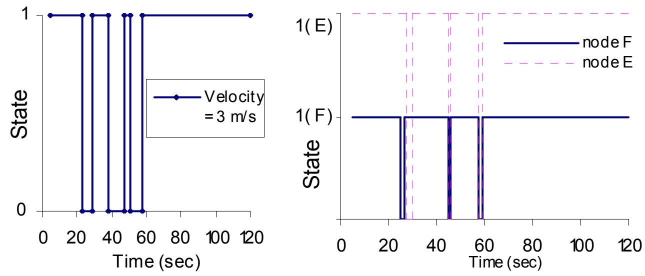


Figure 5 : Scenario under measurement

The duration of disconnected periods depends on mobile node speed. In Figure 6a, we plotted the theoretical distribution of disconnected periods when the velocity of node F is set to 3 m/s. In Figure 6b, we plotted the detected disconnection periods using our passive approach at each end node. Figure 6b shows clearly that the disconnection detector is able to effectively detect disconnection occurrences. As we can notice, the first disconnection period is detected by node E with some delay (≈2 sec). The second and third disconnection periods are roughly detected simultaneously. These observations seem to be natural considering node positions.



(a): Theoretical disconnection distribution

(b): Detected disconnection distribution

Figure 6: Disconnection detection performance

In Figure 7 we plotted the rating factor estimated using PEVOM and VQmon. We varied the mobile node velocity from 1.5 m/s to 6 m/s. As the speed is increased, the disconnection period durations get reduced, resulting in a significant reduction of incurred packet loss rate. In fact, we observed an increase of the rating factor with respect to velocity. Scores obtained using VQmon indicate that the vocal conversation quality has a very poor quality (R<40) making the service unacceptable. However, scores produced using PEVOM indicate that the vocal conversation quality has an acceptable perceptual quality for mobile node velocity greater than 3 m/s (R>60). The observed rating factors are somehow expected. Indeed, VQmon interprets disconnection periods as burst losses in the computation of the rating factor, which degrades dramatically the perceptual quality. However, PEVOM detects disconnection periods and includes appropriately their perceptual effects on mobile users.

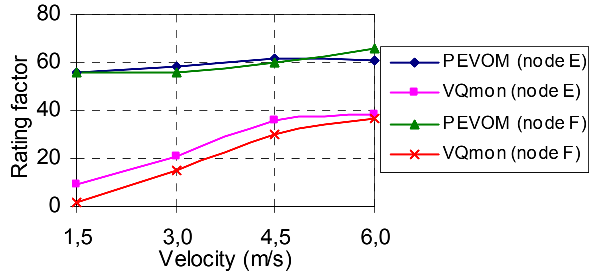


Figure 7: Rating factor according to VQmon and PEVOM

VI. CONCLUSION

In this paper, we have proposed an original instrumental approach denoted PEVOM intended to measure the quality of real-time packet based voice transmission over MANETs. PEVOM detects passively the loss of connectivity and estimates accurately the effect of disconnection during a vocal conversation carried over multi-hop wireless networks. Indeed, connections over mobile ad-hoc networks are intrinsically intermittent. PEVOM captures the time varying impairments by splitting a vocal conversation into connected and disconnected periods. PEVOM ranks appropriately the perceptual quality according to the observed state. Simulation results show that rating factor estimations produced by VQmon are much lower than those obtained by PEVOM. We believe that the approach we adopted in PEVOM is more suitable to rate the quality of voice communication in wireless ad hoc networks. We are in the process of subjectively calibrating the recency effect used in this work to estimate the transmission quality during disconnected periods in order to match scores produced by the subjective and objective rating approaches.

REFERENCES

[1] H. Dong, I. D. Chakares, C.H. Lin, A. Gersho, E. Belding-Royer , U. Madhow, J. D. Gibson, "Speech Coding for Mobile Ad Hoc Networks", Proceedings of the Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, November 2003.

[2] I. Chakeres, H. Dong, E. Belding-Royerz, A. Gersho, J. Gibson, "Allowing Errors in Speech over Wireless LANs", Elsevier Computer Communication Journal, 28, pp. 1643-1657, March 2005.

[3] F. Hammer, P. Reichl, T. Nordstrom, G. Kubin, "Corrupted Speech Data Considered Useful", Proc. First ISCA International Tutorial and Research Workshop on Auditory Quality of Systems, vol. 90 (2004), pp. 1052-1060, Dec 2004.

[4] ITU-T Recommendation P.800, "Methods for Subjective Determination of Transmission Quality", August, 1996.

[5] A.D. Clark, "Modeling the Effects of Burst Packet Loss and Recency on Subjective Voice Quality", In Proceedings of IP Telephony Workshop, Columbia, USA, 2001.

[6] A. E. Conway, "A Passive Method for Monitoring Voice-over-IP Call Quality with ITU-T Objective Speech Quality Measurement Methods", in Proceedings of IEE International Conference Communications, Vol. 4, New York City, NY, pp. 2583-2586, 2002.

[7] R.G Cole, J.H. Rosenbluth, "Voice over IP Performance Monitoring", Computer Communication Review, ACM Sigcomm, 31(2), April 2001.

[8] ITU-T Recommendation G.107, "The E-Model a Computational Model for Use in Transmission Planning", March 2003.

[9] ITU-T Recommendation P.862, "Perceptual Evaluation of Speech Quality (PESQ), An Objective Method for End-to-End Speech Quality Assessment of Narrowband Telephone Networks and Speech Codecs", February, 2001

[10] Miroslaw Narbutt, Mark Davis, "Predicting the impact of application-layer adaptation mechanisms and MAC-layer parameter tuning on speech transmission quality in VoIP systems", Available on-line at http://www.cnri.dit.ie/publications/, visited on January 2008.

[11] L. Sun and E. Ifeachor, "New Methods for Voice Quality Evaluation for IP Networks", In Proceedings of the 18th International Teletraffic Congress (ITC18), Berlin, Germany, September 2003.

[12] J.C. Bolot and A. Vega-Garcia, "Control Mechanisms for Packet Audio in the Internet", In Proceedings of IEEE Infocom'96, April 1996.

[13] ETSI TS 101 329-5 V1.1.1, "Part 5: Quality of Service (QoS) measurement methodologies", 2000-11.

[14] France Telecom R&D, "Study of the relationship between instantaneous and overall subjective speech quality for time-varying quality speech sequence: influence of a recency effect", ITU Study Group 12 Contribution D.139, May 2000.

[15] NS Simulator Homepage, http://www.si.edu/nsnam/ns, visited on 29 Dec 2007.

[16] R.Ramjee, J.Kurose, D.Towsley, and H.Schulzrinne, "Adaptive Playout Mechanisms for Packetized Audio Applications in Wide Area Network", in Proceedings of IEEE INFOCOM, Vol. 2, pp. 680-688, Toronto, Canada, 1994.