Christian Gottron, André König, Matthias Hollick, Sonja Bergsträßer, Tomas Hildebrandt, Ralf Steinmetz:
Quality of Experience of Voice Communication in Large-Scale Mobile Ad Hoc Networks. In: Proceedings of the second IFIP Wireless Days 2009, December 2009.

Quality of Experience of Voice Communication in Large-Scale Mobile Ad Hoc Networks

Christian Gottron*, André König*, Matthias Hollick†, Sonja Bergsträßer*, Tomas Hildebrandt*, Ralf Steinmetz*
* Multimedia Communications Lab (KOM), TU Darmstadt, Germany
{christian.gottron, andre.koenig, sonja.bergstraesser, tomas.hildebrandt, ralf.steinmetz}@kom.tu-darmstadt.de
†Center for Advanced Security Research Darmstadt (CASED), TU Darmstadt, Germany
matthias.hollick@cased.de

***Abstract*—Real-time voice communication is an essential requirement in first responder scenarios. While mobile ad hoc networks (MANET) already prove to be an appropriate communication substrate in small-scale real-world operations, questions regarding scalability limitations remain. In this paper, we identify major factors that affect the quality of experience of voice communication in MANETs. In a series of simulation studies, we show that voice transmission using MANETs is also feasible in large-scale scenarios, if appropriate settings are chosen.**

I. Introduction

Offering features like self-organization and self-healing, MANETs allow for the spontaneous establishment of a network without infrastructure that is able to adapt to a constantly changing topology, which makes them a suitable communication substrate for deployment in first responder scenarios. Yet, due to the multi-hop wireless transmission, MANETs lack the performance and reliability of wired and infrastructure-based networks: this results in a reduced quality of service (QoS) in terms of throughput, delay, loss and jitter. These characteristics of MANETs are challenging for voice communication, which is a basic demand of on-site units in emergency response scenarios.

Although the applicability of MANETs in small-scale real-world scenarios has been shown, questions regarding scalability limitations still remain open. In small-scale scenarios, the aforementioned challenges are negligible due to the usually short route length. In large-scale scenarios, with the size of cities and above, consisting of hundreds of nodes, the increasing route length will strongly affect the quality of a voice transmission.

Common QoS metrics such as throughput, delay, loss, and jitter define a precisely measurable technical description of a networks characteristics. The quality of a voice transmission as it is experienced by the user can only be qualified limitedly in these terms, but can be evaluated by a set of human listeners rating the quality. Clearly, this requires a high effort and is hardly realizable in our context. Instead, tools that are based on human perception models can be utilized.

In this paper, we scrutinize the quality of experience (QoE) of real-time voice streaming in large-scale MANETs. After presenting related work that has motivated our research, we shortly summarize the basics of digital voice transmission and the method we use for the QoE analysis. Our main contribution is the evaluation of the influence of network size and load as well as that of voice codecs and their parameterization on the perceived QoE of voice communication within MANETs.

II. Related Work

In this section we present related work that has motivated our research. We focus, in particular, on projects with comparable application scenarios, as well as on research on voice transmission in general and voice transmission in MANETs.

A MANET-based communication network architecture for large-scale emergency response scenarios is described in [6] and in [7]. Amongst other services, voice communication is one objective of the network design. Yet, to the best of our knowledge, no systematic evaluation of the QoE of the voice communication is performed.

The quality of VoIP streams when transmitted over a MANET is analyzed in [16]. The authors compare the effects of different MANET routing protocols and of node mobility by means of simulation. The metrics used for the evaluation are the technical QoS metrics delay and loss. The scenario is a small-scale ad hoc network consisting of 21 nodes with a low node degree of approximately 4 neighbors per node and an average route length of 2.7 hops. Scalability issues are not considered. A QoE analysis based on human perception models is not performed.

In [15] single-hop 802.11-based ad hoc networks are evaluated subject to the network load in terms of simultaneous voice streams, to node velocity, and to a packet prioritization based on the size of the 802.11 congestion window. The evaluation is performed in terms of the technical QoS metrics loss and jitter. The packet prioritization is evaluated in a small-scale scenario consisting of 40 nodes in an area allowing for direct communication between any two nodes. Node mobility is evaluated in a scenario consisting of 40 nodes with a node degree of approximately 11 neighbors per node and an average route length of 2.2 hops. Neither scalability nor QoE are subjects of the evaluation.

General challenges of voice communication in MANETs are identified in [3]. Aspects of MAC and routing protocols for MANETs are discussed. Different approaches of speech coding are reviewed with respect to their applicability in MANETs. A QoE analysis is not part of the work. Scalability issues are not considered.

Besides 802.11-based ad hoc networks, Terrestrial Trunked Radio - Direct Mode Operation (TETRA DMO) [4] offers a solution for wireless ad hoc communication. However, this standard is (to the best of our knowledge) not tailored to the extensive multihop communication we consider in this work.

III. Codecs, Protocols & QoE Analysis

In this section we provide background information on the relevant details of voice coding, voice transmission, and QoE analysis as a basis for the remainder of this paper.

A. Voice Coding

The first step for voice communication in MANETs is the application of a voice codec for conversion from an analog signal to a digital signal. Voice codecs can be categorized in (1) vocoders, (2) waveform codecs, and (3) hybrid codecs. These types show fundamental differences with respect to their architecture and their mode of operation. As a result, the bandwidth required and the speech quality achieved differ strongly. Vocoders require a bandwidth of only few (single-digit) kbit/s. In consequence, the speech quality is poor. Waveform codecs, on the other hand, can produce an excellent speech quality, but have bandwidth requirements, of an order of magnitude, higher than that of vocoders. Hybrid codecs can be classified somewhere between vocoders and waveform codecs, regarding the bandwidth requirements and speech quality achieved. For further information, we refer to [5]. Due to the inherently poor speech quality of vocoders, we focus on waveform codecs and hybrid codecs with G.711 [11] and Speex [14] as particular representations in this work.

G.711 is a well known waveform codec, which is based on pulse code modulation. The uncompressed and high quality codec requires a bandwidth of 64 kbit/s. An analog signal is sampled at a rate of 8 kHz and quantized on an 8 bit logarithmic scale resulting in a bandwidth requirement of 64 kbit/s. One of the most prominent fields of application of G.711 is the ISDN telephone system. Speex is a hybrid codec based on the code-excited linear prediction algorithm [9]. A parameterization of Speex is possible regarding sample rate, speech quality, and complexity of the encoding algorithm. As a result, the bandwidth required varies from 2.15 kbit/s to 44 kbit/s.

B. Voice Transmission in MANETs

Transmission of encoded voice via a MANET can be seen as a two-step process. In the first step, a route between sender and receiver is established. This can be done either proactively or reactively. In proactive routing protocols, each node constantly keeps track on how other nodes of the network can be reached. For this, corresponding topology information is exchanged periodically. Reactive routing protocols determine a route not before it is required. Which class of routing protocol should be deployed, depends on the MANET application area. Proactive protocols produce a constant base-load of the network to keep the routing tables of all nodes up to date. In our scenario, we assume high node mobility combined with a relatively sparse network. For this reason, we base our study on the reactive ad hoc on-demand distance vector (AODV) routing protocol as described in [8]. Here, to establish a route between a source and a destination, a route request message is sent as broadcast by the source. The request is forwarded by intermediate nodes until it reaches the destination. Each intermediate node keeps track on the predecessor from which the request was received, thus establishing a reverse route. When the request reaches the destination, a route reply message is generated and sent along the reverse route to the source. Upon receiving the route reply at the source, the route is established.

After a route is established, the second step conducts the transmission of the encoded voice. For transmission, we deploy the real-time transport protocol (RTP), as described in [10]. The primary task of RTP is to ensure correct packet ordering. Thus, RTP packets contain a sequence number and a timestamp. The combination of RTP and UDP allows for a reliable transmission of multimedia data with a minimum of protocol related delay and overhead.

In order to transmit voice data over packet-oriented networks, the codec generates data frames of a specific length. For G.711, this interval is variable and part of our evaluation. For Speex, the interval is fixed to $20ms$ due to the compression algorithm, which is part of the coding process.

The variance of the delay (jitter) of subsequent packets demands usage of a jitter buffer. With this, the first packet of a voice stream is buffered for a certain amount of time before it is handed to the codec in order to compensate the jitter of the following packets. The size of the jitter buffer defines the maximum variance of the delay accepted. Packets that exceed this maximum are dropped.

C. QoE Analysis

The aim of this paper is to evaluate the QoE of voice transmissions. The QoE can be derived from the voice quality as experienced at the receiver and from the transmission delay.

The International Telecommunication Union (ITU) defines the voice quality by the Mean Opinion Score (MOS) [13]. The MOS specifies five different categories of voice quality, from 1 (worst) up to 5 (best). In the best case, the G.711 codec has a 4.4 MOS while the quality of Speex lies between 2 and 4, depending on the codec settings.

The MOS can be determined subjectively, which requires a sufficiently large number of listeners to rate the quality. Thus, this subjective determination of the voice quality is not feasible in our context. An objective evaluation of the MOS can be performed automatically, based on human perception models. For this, the ITU proposed the Perceptual Evaluation of Speech Quality (PESQ) algorithm [12]. This algorithm is designed to determine the quality of a narrowband voice stream. PESQ compares the original voice stream with the one received. The correlation between a subjective determination of the voice quality and the corresponding PESQ result is 0.935.

IV. QoE Analysis in MANETs

In this section, we present the analysis of the QoE of voice communication in MANETs. After describing the experimen-

tal design, we evaluate the effects of (1) the network size, (2) the frame size, (3) the network load and (4) the voice codec.

A. Experimental Design

For the simulation studies, we used an extended version of the JiST/SWANS simulation tool [2]. We added a traffic model for voice communication in the first responder scenarios. The default simulation settings are given below. We used these if not specified otherwise.

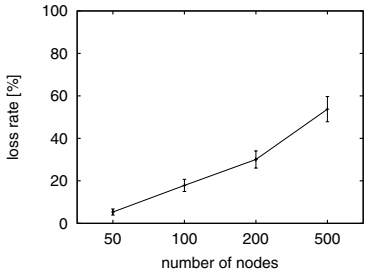
Transmission range	$250m$
Average neighbors per node	8
Voice streams	16 different recordings as recommended by the ITU [12]
Codec	G.711
Network load	5 simultaneous voice streams
Frame size	$20ms$
Jitter buffer	$100ms$
Average processing delay per hop	$1ms$
Initial placement	Random with uniform geographical distribution
Mobility	Random waypoint with continuous movement and minimum and maximum speed of $1m/s$ and $2m/s$
Communication model	Random selection of sender and receiver.
Simulated time	$600s$ per setup

The default network parameters and simulation settings are chosen such that we obtain a moderately loaded network. While transmission range may vary due to the environment, we chose a fixed value of $250m$ in an environment without obstacles to reduce the influences on voice quality for the scope of this paper. Further, the density of the network was chosen such that a connected (unpartitioned) network is typically achieved. The random waypoint mobility model can be considered as a worst case scenario with respect to the predictability of node movement. The simulated time per factor set was split up into multiple simulation runs to reduce any unwanted side-effects of the random waypoint model. While more realistic mobility models for first-responder scenarios were proposed [1], for our simulations we chose the random waypoint model as an unpredictable worst case scenario.

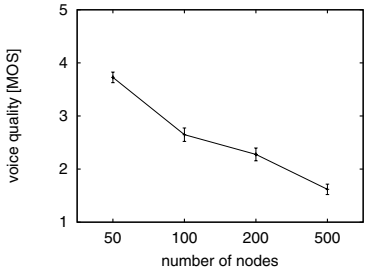
The protocols used in the network are as follows.

Physical layer	Segmental calculation of the signal power and the SNR
MAC Layer	IEEE 802.11 DCF with a transmission rate of $11Mbit/s$
Network Layer	IPv4 with AODV routing service and buffers up to 127 packets
Transport Layer	UDP with RTP

In the evaluation, we focus on the metrics (1) voice quality, (2) packet loss, and (3) transmission delay. The voice quality was determined using the PESQ evaluation tool available at [12]. The results were converted to the MOS scale to be compared to related work. The delay is measured in two different ways. The overall end-to-end delay is calculated as the average of the delay of all packets, including the delay that results from the route discovery process of the reactive AODV routing. While the routing delay is a major part of the delay, it only affects the very beginning of a voice stream (like the waiting time after dialing in a traditional telephone call). Thus, we also present the transmission delay of the network, which does not include the routing delay.



(a) Packet loss for the variation of the network size



(b) Voice Quality for the variation of the network size

Fig. 1. Variation of the network size

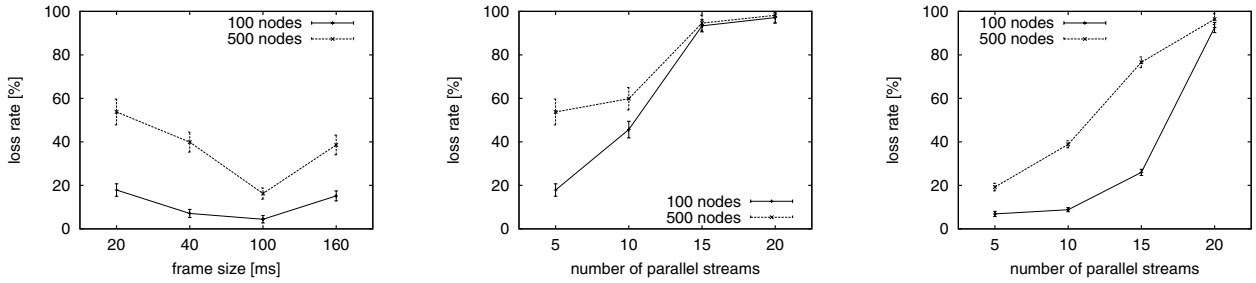
B. Effects of the Network Size

Firstly, we illustrate the dependency between the network size and the voice quality. For this, we evaluated a small (50 nodes in a $1050m \cdot 1050m$ area), medium (100 nodes in a $1500m \cdot 1500m$ area) and large (500 nodes in a $3300m \cdot 3300m$ area) scenario. The area was chosen such that the average number of neighbors per node (i.e. the probability for a connected network) is not affected.

With an increasing network size, in terms of nodes, we observed an increased route length and with this, an increased number of packet collisions and mobility-induced route breaks. As a result, the packet loss rate also increased, as shown in

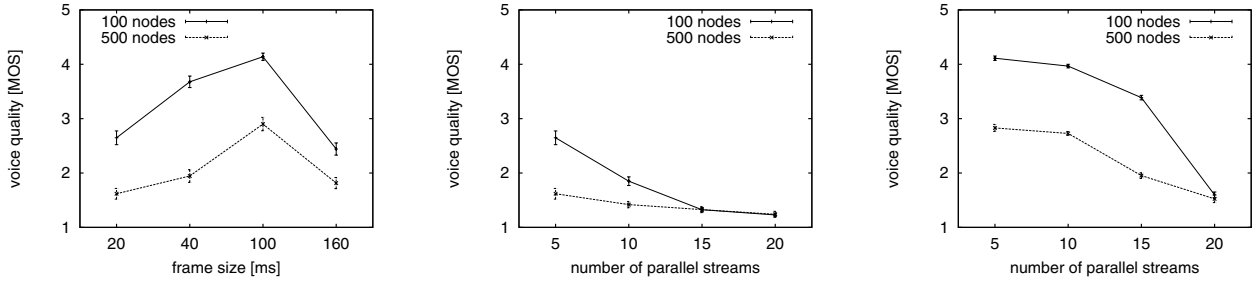
TABLE I
ROUTE LENGTH, DELAY, AND JITTER FOR THE VARIATION OF THE NETWORK SIZE

Nodes	Route length	Overall delay	Trans. delay	Jitter
50	2.53	76.5 ms	24.9 ms	4.5 ms
100	3.50	160 ms	32.8 ms	11.5 ms
500	7.40	584 ms	73.2 ms	48 ms



(a) Variation of the frame size (b) Variation of the network load (20ms frames) (c) Variation of the network load (100ms frames)

Fig. 2. Packet loss for the variation of the frame size and the number of parallel streams



(a) Variation of the frame size (b) Variation of the network load (20ms frames) (c) Variation of the network load (100ms frames)

Fig. 3. Voice Quality for the variation of the frame size and the number of parallel streams

Figure 1(a). Due to this, the voice quality rapidly deteriorates as the network size increases. As shown in Figure 1(b), a good voice quality (3.7 MOS) was achieved in the small-scale scenario. The medium-scale and large-scale scenarios showed a heavily reduced voice quality due to the increased loss rate.

The delay and jitter were affected in a similar way as shown in Table I. The overall end-to-end delay of the large-scale scenario was above the $500ms$ limit that was considered acceptable for voice communication: yet, the transmission delay of the network was acceptable for voice transmission. The small and medium-scale scenarios showed both an acceptable overall delay and transmission delay. For all scenarios, the jitter was between $4.5ms$ and $48ms$ and could, thus, be completely compensated by the jitter buffer.

C. Effects of the Frame Size

The G.711 codec enables a flexible selection of the frame size, as no compression is performed. For the following, we varied this parameter from $20ms$ to $160ms$ for both the medium (100 nodes) and large (500 nodes) scenarios.

Up to the size of $100ms$, an increased frame size positively affected the loss rate, as can be seen in Figure 2(a). While the loss rate in the large-scale scenario with $20ms$ frames was about 54%, a frame size of $100ms$ reduced the loss by more than 35%. Due to the modified loss rate, the MOS rate was improved by 1.4 points in the large-scale scenario as shown in Figure 3(a). A similar effect was observed in the medium-scale scenario. Altogether, adapting the frame size resulted in a reduced loss rate and thus, in an increased voice quality. On the downside, an increased frame size caused higher, yet tolerable overall and transmission delays, as shown in Table II.

D. Effects of the Network Load

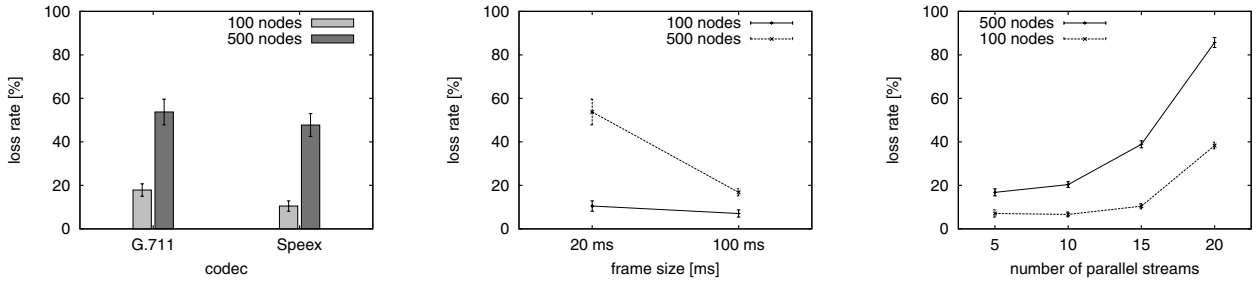
As shown in the previous section, reducing the network load by increasing the frame size resulted in a reasonable QoE of voice communication in all scenarios. Thus, we directly analyzed how the network load, in terms of parallel voice streams, affected the QoE. Starting from the 5 parallel streams of the previous experiments, we increased the load up to 20 parallel streams. Again, we focused on the medium-scale and on the large-scale scenario. For each scenario, we determined

TABLE II
DELAY FOR THE VARIATION OF THE FRAME SIZE

Frame length	100 nodes		500 nodes	
	Overall delay	Trans. delay	Overall delay	Trans. delay
20	160 ms	32.84 ms	584 ms	73.2 ms
40	163 ms	89.3 ms	686 ms	82.1 ms
100	269 ms	109.9 ms	726 ms	128.9 ms
160	548 ms	171.4 ms	1102 ms	224.9 ms

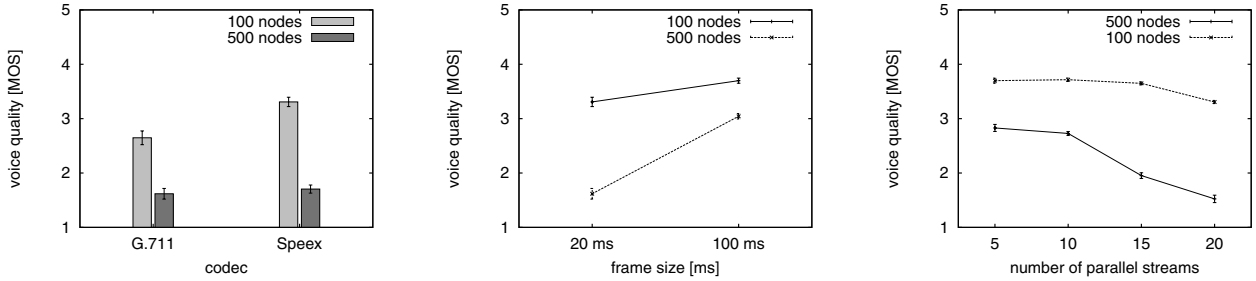
TABLE III
TRANSMISSION DELAY FOR THE VARIATION OF THE NETWORK LOAD

Streams	100 nodes		500 nodes	
	20 ms	100 ms	20 ms	100 ms
5	32.8 ms	109.91 ms	73.2 ms	158.4 ms
10	87.6 ms	113.01 ms	113.4 ms	191.7 ms
15	151.7 ms	231.43 ms	135.2 ms	231.4 ms
20	76.60 ms	235.60 ms	127.2 ms	235.6 ms



(a) Variation of the network size (b) Variation of the frame size (c) Variation of the network load

Fig. 4. Packet loss for varied network parameters with the Speex codec



(a) Variation of the network size (b) Variation of the frame size (c) Variation of the network load

Fig. 5. Voice quality for varied network parameters with the Speex codec

the influence of the network load for both the standard frame size of $20ms$ of G.711 and the $100ms$ optimum.

Increasing the network load led to an increased packet loss, due to collisions, in both scenarios. The effect was amplified for a small frame size of $20ms$, while $100ms$ frames provided low packet loss even in scenarios with increased network load, as shown in Figures 2(b), 2(c), 3(b) and 3(c). The large-scale scenario still showed a tolerable loss rate and a fair voice quality for up to 10 parallel streams and a frame size of $100ms$. In the medium-scale setup, with $100ms$ frames, the network load could be increased up to 15 parallel streams, while maintaining a feasible packet loss and a good voice quality. What stands out here was that the packet loss caused by the network for 15 parallel streams in the large-scale scenario was about the same as the loss for 5 parallel streams in the medium-scale scenario. Yet, the voice quality for these settings differed strongly, whereas within the experiments regarding the frame size presented in the previous section, a comparable loss rate for the medium-scale and large-scale scenarios resulted in a comparable voice quality. This could be explained by the fact that the jitter in the large-scale scenario exceeds the jitter-buffer of $100ms$. Thus, in addition to the packet loss that is caused by the network itself, packets have to be discarded due to the jitter, which results in the deterioration of voice quality.

The overall delay and the transmission delay for the variation of the network load are shown in Tables IV and III. Both increased along with the network load. Yet, the transmission delay was tolerable for all levels of network load.

E. Effects of the Voice Codec

Up to now, we used the waveform codec G.711 for voice coding. In this section, we compare G.711 to the hybrid codec Speex. In contrast to G.711, Speex uses a compression algorithm, thus reducing bandwidth requirements and also reducing the maximum voice quality achievable. Speex allows for adaption of this compression algorithm in order to further reduce the bandwidth requirement. Since our goal is to optimize the voice quality, we used the highest quality setting of Speex. With this, the bandwidth required is $25.6kbit/s$ which is still less than half the bandwidth requirement of G.711 ($64kbit/s$). Due to the compression algorithm, Speex

TABLE IV
OVERALL DELAY FOR THE VARIATION OF THE NETWORK LOAD

Streams	100 nodes		500 nodes	
	20 ms	100 ms	20 ms	100 ms
5	160 ms	189 ms	583 ms	652 ms
10	285 ms	220 ms	673 ms	765 ms
15	473 ms	259 ms	750 ms	1009 ms
20	490 ms	578 ms	782 ms	1015 ms

TABLE V
TRANSMISSION DELAY FOR THE VARIATION OF THE NETWORK SIZE AND THE NETWORK LOAD WITH THE SPEEX CODEC

Streams	100 nodes	500 nodes
5	106.7 ms	137.0 ms
10	109.2 ms	142.8 ms
15	110.9 ms	167.0 ms
20	117.0 ms	188.0 ms

requires a fixed frame size of $20ms$. For comparison to the increased frame size of G.711 we thus sent multiple frames per packet when we used the Speex codec. We compared G.711 and Speex for frame sizes / multiple frames per packet of $20ms$ and $100ms$. We also compared both codecs subject to the network load for the $100ms$ frames.

Since the maximum voice quality achievable is higher for G.711 than for Speex, the voice quality of G.711 exceeded that of Speex for the medium-scale scenario as shown in Figure 5(a). Yet, in the large-scale scenario, the reduced bandwidth requirement of Speex led to a better voice quality compared to G.711, due to a lower packet loss as shown in Figure 4(a).

The qualitative results for a variation of the network's size and workload when using Speex were comparable to the results for G.711, as presented in the previous sections. For both codecs, the increased packet loss and thus, the reduced voice quality could be compensated by increasing the frame size or the number of frames per packet, respectively, as shown in Figures 5(b) and 4(b).

Also, due to the low bandwidth requirement of Speex, a higher network load could be applied. As shown in Figure 5(c), the voice quality was always between fair and good for the medium-scale scenario. For the large-scale scenario, a fair voice quality could be maintained for a load of up to 15 parallel streams. The packet loss, as shown in Figure 4(c), supports this result. Again, it stands out that a comparable packet loss did not lead to a comparable voice quality in the different scenarios. As for the results presented in the previous section, we ascribe this to a jitter that exceeded the $100ms$ jitter buffer. The effect was also slightly amplified by the fact that sending multiple frames per packet resulted in a more bulky packet loss compared to the statistically distributed loss, which was observed if one frame was sent per packet. With respect to the resulting voice quality, Speex was affected to a greater extent by this than G.711, due to the different modes of operation of the codecs.

The transmission delay, shown in Table V, when using Speex was slightly lower than for G.711: this is due to the reduced network load leading to a lower congestion related delay.

V. Conclusion

In this paper, we analyzed major factors influencing the QoE of voice communication in MANETs. We focused on: (1) the network size in terms of nodes and geographical expanse, (2) the length of voice frames, (3) the network load in terms of parallel voice streams and (4) the voice codec deployed.

Our results show that increasing the network's size and workload leads to a reduced voice quality due to an increased packet loss rate. Since the network's size and load are specified by system requirements and user behavior, these parameters cannot be adapted or limited unrestrictedly. To improve the voice quality in large-scale and/or heavily loaded networks, measures such as changing the codec deployed and increasing the frame size can be taken. The choice of the codec directly affects the maximum achievable voice quality. Due to the individual bandwidth requirements of different codecs, the network load is also affected. Increasing the frame size or the number of frames per packet reduces network load and protocol related overhead. To a certain extent, an adaptation of the codec and the frame size can thus counterbalance possible negative effects of large-scale and heavily loaded networks, allowing for voice communication with a reasonable quality of experience, in these scenarios. Furthermore, our studies show that the statistical packet loss can only be used as an indicator for the QoE. While packet loss always results in a weaker QoE, scenarios presenting the same packet loss rate may differ greatly in terms of voice quality.

References

[1] N. Aschenbruck, P. Martini, and M. Gerharz. Human Mobility in MANET Disaster Area Simulation. - A Realistic Approach. In *Proceedings of the 29nd Annual IEEE International Conference on Local Computer Networks (LCN 04)*, 2004.

[2] R. Barr. *An efficient, unifying approach to simulation using virtual machines*. PhD thesis, Cornell University, 2004.

[3] H. Dong, I. Chakares, C. Lin, A. Gersho, E. Belding-Royer, U. Madhow, and J. Gibson. Speech coding for mobile ad hoc networks. In *Proceedings of the Thirty-Seventh Asilomar Conference on Signals, Systems and Computers (ACSSC 03)*, 2003.

[4] ETS 300 396-1. *Terrestrial Trunked Radio (TETRA); Technical requirements for Direct Mode Operation (DMO); Part 1: General network design*. ETSI, 1998.

[5] L. Hanzo, C. Somerville, and J. Woodard. *Voice and Audio Compression for Wireless Communications*. John Wiley & Sons, Inc., second edition, 2007.

[6] K. Kanchanasut, A. Tunpan, M. Awal, D. Das, T. Wongsaardsakul, and Y. Tsuchimoto. A Multimedia Communication System for Collaborative Emergency Response Operation in Disaster-affected Areas. Technical report, Internet Education and Research Laboratory (intERLab), Asian Institute of Technology (AIT), 2007.

[7] W. Lu, E. C. Peh, W. Seah, and Y. Ge. Communications Support for Disaster Recovery Operations using Hybrid Mobile Ad-Hoc Networks. In *Proceedings of the 32rd Annual IEEE International Conference on Local Computer Networks (LCN 07)*, 2007.

[8] C. E. Perkins, E. M. Belding-Royer, and S. R. Das. Ad hoc On-Demand Distance Vector (AODV) Routing. *IETF RFC 3561*, 2003.

[9] M. R. Schroeder and B. S. Atal. Code-excited linear prediction(CELP): High-quality speech at very low bit rates. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 85)*, 1985.

[10] H. Schulzrinne and et al. RTP: A Transport Protocol for Real-Time Applications. *IETF RFC 3550*, 2003.

[11] SERIES G. *G.711: Pulse Code Modulation (PCM) of Voice Frequencies*. ITU-T, 1988.

[12] SERIES P. *P.862: Perceptual Evaluation of Speech Quality (PESQ)*. ITU-T, 2001.

[13] SERIES P. *P.800.1 Mean Opinion Score (MOS) terminology*. ITU-T, 2006.

[14] J. Valin. *The Speex Codec Manual Version 1.2 Beta 3*. Xiph.org Foundation, December 2007.

[15] P. Velloso, M. Rubinstein, and O. Duarte. Analyzing voice transmission capacity on ad hoc networks. In *Proceedings of International Conference on Communication Technology (ICCT 03)*, 2003.

[16] H. Zhang, J. Homer, G. Einicke, and K. Kubik. Performance Comparison and Analysis of Voice Communication over Ad Hoc Network. In *Proceedings of the 1st Australian Conference on Wireless Broadband and Ultra Wideband Communications (AusWireless 06)*, 2006.