Perceptual Voice Communications in IEEE802.15.4 Networks for the Emergency Management Support

Matteo Petracca*‡, Marco Ghibaudi†‡, Riccardo Pelliccia†‡, and Paolo Pagano*‡
*National Laboratory of Photonic Networks, Consorzio Nazionale Interuniversitario per le Telecomunicazioni, Pisa, Italy
†Scuola Superiore Sant'Anna Research Unit, Consorzio Nazionale Interuniversitario per le Telecomunicazioni, Pisa, Italy
‡Real-Time Systems Laboratory, Scuola Superiore Sant'Anna, Pisa, Italy
Email: [matteo.petracca|paolo.pagano]@cnit.it, [m.ghibaudi|r.pelliccia]@sssup.it

Abstract*—Voice communication in IEEE802.15.4 networks is an attractive application for the emergency management support when a disaster occurs. Although this application can be of extreme utility in the immediate consequence of the disaster, it must take into account the bandwidth limitations of the IEEE802.15.4 standard, which allows the simultaneous transmission of a small number of voice streams. When more voice streams must be transmitted through the network, a bandwidth reduction of each speech flow must be performed. In this paper we first present an algorithm for the perceptual selection of voice data aiming at reducing the speech flow bandwidth while preserving as much as possible the end-to-end speech quality, then we propose a voice data protection technique based on speech perceptual importance and able to preserve speech quality against packet losses. The perceptual selection algorithm can reach a 30.8% reduction in bandwidth occupancy, with respect to a full rate voice communication, still maintaining an end-to-end speech quality between good and fair according to the Mean Opinion Score scale defined by the International Telecommunications Union. The protection technique, jointly adopted with the perceptual selection, can reach better end-to-end speech quality values with respect to a full rate voice communication while requiring a lower amount of transmission bandwidth.

***Keywords*-Perceptual voice communications; Perceptual voice protection; IEEE802.15.4.**

I. Introduction

In recent years Wireless Sensor Networks (WSNs) have experienced a rapid growth both in academia and industry. If from one hand the research community went ahead in developing new technological solutions going towards the full accomplishment of the Internet of Things (IoT) [1], [2] concept, on the other hand private stakeholders started to benefit of such kind of technology proposing WSN systems for creating smart ambients focusing on military, energy-efficient buildings and intelligent transportation scenarios.

For the next future years it is possible to envision all cities covered by wireless sensors providing a pervasive monitoring infrastructure in support of human activities. Pervasive network of sensors will be used to detect in real-time the car traffic conditions in a new generation of Intelligent Transportation Systems (ITSs) [3]–[6], while WSNs installed in apartments and buildings will provide and guarantee comfort and security to the living people [7]–[10].

In the above mentioned scenarios, wireless sensor devices can be pervasively deployed in the environment to ensure a capillary acquisition of information, thus enabling the possibility of creating new effective emergency support applications [11]. In case of a natural disaster, e.g. an earthquake, buildings and roads can collapse entrapping people, while at the same time telecommunication infrastructures can be completely unusable due to lack of electricity or equipment failures, thus delaying rescue operations. Wireless sensor networks are composed by tiny mote devices which have more chance than regular telecommunication infrastructures to survive to the effects of a natural disaster, while still providing a communication service. In a WSN, in fact, the devices are battery powered (no dependence on the grid), and although some of them can experience failures, their pervasive deployment in the environment can still guarantee reliable communications. As a consequence, it is possible to think wireless sensor network technology as an effective solution for providing support in emergency situations [12]. When a natural disaster occurs the WSNs functions can be changed at run-time in order to provide essential services, such as voice communications, to the rescuers. The possibility of enabling voice communications among rescuers in a short time is useful to give them the possibility of better coordinate their action in order to save the greatest number of people. Moreover, in case of entrapped people close to microphone equipped devices, a voice communication can be used to better assist them during the rescue operations.

Voice communications for emergency support in a WSNs scenario have been considered in several works published over the last years. In [13] the performance of a voice communication system to support workers in a coal mine is analyzed by means of a real implementation in the National Institute for Occupational Safety and Health (NIOSH) experimental coal mine in Pennsylvania, while in [14] and [15] the problem of enabling end-to-end speech quality is addressed by means of voice packets duplication and perceptual based voice packets

DOI: 10.1109/BCFIC.2012.6217991

protection. In all the above mentioned works the whole voice stream is supposed to be sent through the network and its reliability ensured by means of additional information. The final result of this approach is an improvement of the speech quality at the cost of a higher bandwidth occupancy. In wireless sensor networks based on the IEEE802.15.4 standard [16] the available bandwidth is a very limited resource, 250 Kbit/s at the physical (PHY) layer with even lower values at the medium access control (MAC) layer (e.g., 163 Kbit/s in case of unslotted based transmissions [17] and slightly bigger values in case of slotted based transmissions [18]). Considering the use of the low-complexity speech compression standard ITU-T G.711 [19] (0.01 million instructions per second in complexity [20]) for enabling voice communications in a WSN scenario the required bandwidth at application layer is equal to 64 Kbit/s which increases to a value of 82.4 Kbit/s at MAC layer assuming a packet size able to transport 10 ms of voice and regular IEEE802.15.4 header fields. With the bandwidth occupancy value reported above only one voice flow can be transmitted in case of IEEE802.15.4 unslotted based communications and no more than two in case of slotted based transmissions. When multiple streams must be transmitted through the network the bandwidth of each of them must be reduced while preserving the end-to-end speech quality as much as possible.

In this paper we present a perceptual based solution for reducing the bandwidth occupancy of a single voice communication stream in IEEE802.15.4 networks, thus providing the possibility of enabling more simultaneous speech flows within the network. The main benefit of the proposed solution is evaluated in terms of bandwidth occupancy reduction while considering its impact on end-to-end speech quality quantified in terms of mean opinion score (MOS) [21] values evaluated through the perceptual evaluation speech quality (PESQ) [22] method. Moreover, we present a perceptual based data protection technique compliant with the IEEE802.15.4 standard and able to recover speech data losses in error-prone channels.

The rest of the paper is organized as follows: in Section II we present the perceptual based solution for reducing the voice flow bandwidth occupancy and its impact on the end-to-end speech quality. In Section III we present the developed perceptual based data protection technique while providing its performance in a simulated multi-hop communication scenario in which real loss traces have been adopted. Conclusions follow in Section IV.

II. Perceptual based voice data selection

A. Voice communications in IEEE802.15.4 networks and perceptual based bandwidth occupancy reduction

In voice communications over packet based networks the speech signal is usually transmitted in a compressed format in order to reduce the required transmission bandwidth. The compression effectiveness reached from a particular compression standard is proportional to its complexity in terms of memory occupancy and required computational capacity. The lower is the reached bit rate, the higher is the complexity of a particular compression algorithm. In a WSN context the sensor devices are usually equipped with low-power and low-end microcontrollers where the available memory is a very scarce resource. Due to the aforementioned limitations in hardware devices, the selected compression standard adopted in this work is the ITU-T G.711 which is characterized by only 0.01 million instructions per second (MIPS) of complexity and 64 Kbit/s of bit rate. Concerning the dimension of each compressed speech packet to be sent through the network, this has been chosen taking into account the maximum available data payload provided by the IEEE802.15.4 standard. More in particular, in this work the packet dimension has been imposed equal to 80 bytes (10 ms of voice), as popular VoIP speech compression standards, against a maximum available MAC data payload of 102 bytes (maximum packet size is 127 bytes). The choice of not using the whole available payload to send voice data from one hand could results in a non optimal bandwidth utilization, while on the other hand it enables the possibility to use the residual bytes for data protection techniques able to recover data losses in error-prone channels.

According to the selected compression standard and voice packet dimension, considering a header size equal to 21 bytes, the required bandwidth at MAC layer for each voice flow is equal to 82.4 Kbit/s (final packet dimension is given by 21 bytes for header plus 80 bytes for data plus 2 bytes for the IEEE802.15.4 frame check sequence), as briefly introduced in Section I. The above reported necessary bandwidth has been evaluated under the hypothesis that all voice packets are sent through the network, and hence that all packets have the same importance. In order to reduce the voice stream bandwidth occupancy only the most perceptual important packets, or segment of them, can be sent. The perceptual importance of a speech packet can be expressed in terms of the distortion that would be introduced by its loss, or non-transmission, into the received speech flow [23]. A distortion measure is aiming at comparing the original speech flow (before the compression) with its decompressed and reconstructed version where data losses have been overtaken by means of concealment techniques. Common distortion measures are reported in [24]. In this work the adopted distortion measure is the signal energy while the adopted concealment is the silent insertion due to its low-complexity requirements. Under these conditions the signal energy of a speech segment to be sent is exactly the distortion that would be introduced by its loss (energy of the inserted silent speech segment equal to zero). The higher is the speech segment energy, the higher is the distortion introduced by its loss or non-transmission through the network.

In the proposed approach the perceptual importance of a speech flow is not evaluated per packet, thus selecting if a packet has to be sent or discarded, but it is evaluated for packet segments. Before compression the amount of speech samples to be compressed in one packet are divided in an imposed number of segments, N_{seg}, and for each one of them the energy is evaluated. Among all the evaluated energy values, E_i, the maximum value is identified, E_{max}, and only the segments in which E_i/E_{max} in percentage is bigger than the imposed selection threshold S_{tr} are compressed and sent through the network. In case E_{max} is equal to zero the full packet is discarded (absence of active voice). As result of the proposed perceptual selection algorithm, at least one active voice segment of a packet (the most important) is sent through the network. To be noted that the full perceptual evaluation algorithm does not require to store previous speech packets, thus avoiding to use additional device memory, a very scarce resource in WSN motes. The pseudocode of the described algorithm can be summarized as follows:

Set N_{seg} *and* S_{tr} *parameters*
for Each uncompressed speech packet **do**
  for $i = 1 \to N_{seg}$ **do**
    Evaluate E_i
  end for
  Search for E_{max}
  if $E_{max} = 0$ **then**
    Discard all segments
  else
    Compress and select for transmission
    all segments in which $[100 \cdot (E_i/E_{max})] \geq S_{tr}$
  end if
end for

The impact of the proposed perceptual selection algorithm on end-to-end speech quality has been evaluated as a function of S_{tr} for N_{seg} values equal to 2, 4, 8 and 16 simulating the transmission of 16 speech traces, each one 3 minutes long, belonging to the NTT Multi-lingual Speech Database [25]. In the performed simulations no network losses have been considered to evaluate the performance of the developed algorithm as a function of its parameters, while the speech quality has been evaluated with the PESQ method, and obtained values averaged and mapped into the much more common MOS scale (Excellent=5, Good=4, Fair=3, Poor=2, Bad=1) according to [26]. The results of the performed analysis are depicted in Figure 1. In the same figure the MOS behavior when the speech segments to be sent are selected in a random way is also reported. In the random selection evaluation process several simulation runs have been performed selecting the segment to be sent according to a uniform distribution and imposing for each value of the threshold an equal amount of speech segments with respect to the perceptual selection. Regarding the results obtained with the perceptual selection it is possible to see how for S_{tr} equal to zero (only silence packets are discarded) MOS is equal to 4.38, which is the same value obtained sending the whole speech flow. Increasing the threshold, a reduction in transmitted voice data information is accepted and in general MOS decreases until reaching a bad quality. The different curves behavior for the various N_{seg} values is mainly due to the different amount of information transmitted and will be better investigated in Section II-B. Considering the random selection results a poor speech quality is reached for a threshold value equal to zero. In this case, in fact, both silent and active speech packets can be discarded, unlike the perceptual selection where only silent packets are discarded, thus suddenly lowering the end-to-end speech quality. Fig. 1 shows that in any case a perceptual selection outperforms a random selection.

B. Bandwidth occupancy and speech quality analysis

If from one hand the increase of the threshold value affects in a negative way the received speech quality, on the other hand it guarantees to save useful transmission bandwidth. In Fig. 2 the behavior of the bandwidth occupancy at MAC layer for a single speech flow is depicted as a function of the selective threshold for all the adopted values of N_{seg}. The bandwidth occupancy behavior is the same for the perceptual and random selection due to the imposed constraints in evaluating the random selection performance. For S_{tr} equal to zero the required bandwidth is maximum and common

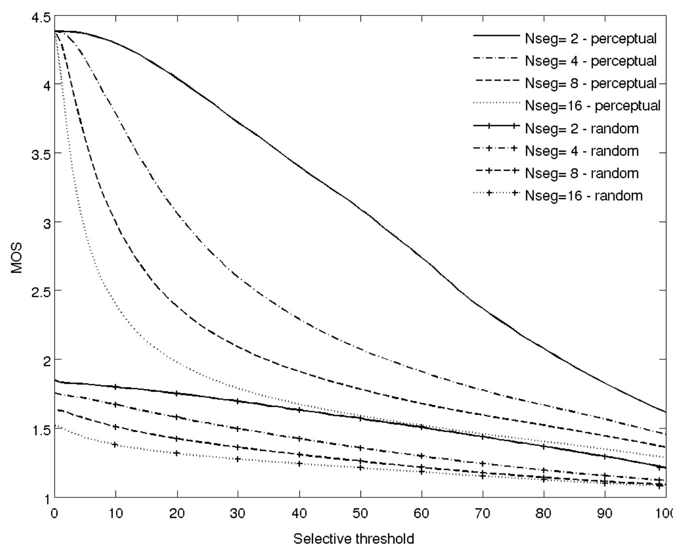

Fig. 1. MOS as a function of the selective threshold.

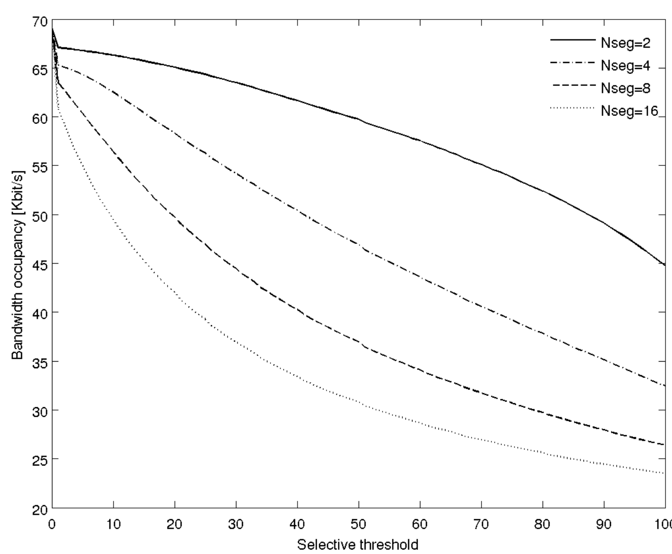

Fig. 2. Bandwidth occupancy as a function of the selective threshold.

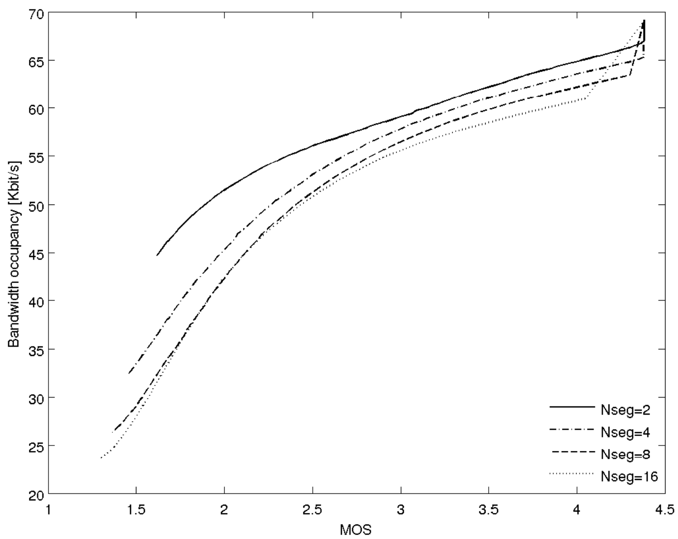

Fig. 3. Bandwidth occupancy as a function of the MOS values.

among all the considered values of N_{seg}. In case of bigger values of S_{tr}, having fixed N_{seg}, the required bandwidth decreases, explaining again the reduction of speech quality due to the decrease of available data at the receiver. Considering a fixed value of S_{tr} the required bandwidth decreases when N_{seg} increases. This behavior is due by the finer granularity wherewith the most perceptual important parts of a packet are selected which gives the possibility to discard more segments.

An interesting analysis of the obtained results is depicted in Fig. 3, where the required bandwidth for a voice communication is plotted as a function of the obtained speech quality. Apart from the point in which MOS is maximum and all the curves converge to the maximum bandwidth value, as it happens in Fig. 1 and 2 when S_{tr} is equal to zero, the same speech quality is reached with a lower bandwidth occupancy increasing the N_{seg} value. If we consider a speech quality between fair and good, MOS equal to 3.5, this is reached with a required bandwidth of 58.2 Kbit/s for N_{seg} equal to 16 and with bigger bandwidth values for lower values of N_{seg}.

The graph in Fig. 3 well summarizes the aim of the proposed algorithm: accepting to reduce the final end-to-end speech quality according to a perceptual criteria in order to save network bandwidth for enabling multiple voice communications. According to the data reported above, in which the accepted MOS is imposed equal to 3.5, the highest reached bandwidth reduction is equal to 30.8% at the cost of a MOS reduction equal to 0.88. A substantial reduction in bandwidth occupancy is reached while still maintaining an end-to-end speech quality between good and fair (between perceptible but not annoying and slightly annoying distortion values). When lower speech quality values can be accepted at the receiver side bigger bandwidth reduction values can be reached and more voice flows can be enabled.

III. Perceptual based voice data protection

A. Data protection strategy

When voice data information are sent through error-prone channels, as it happens in IEEE802.15.4 based communications, the occurred packet loss negatively affects the end-to-end speech quality [27]. A solution for recovering useful data information consists in adopting data protection techniques aiming at reaching the best trade off between additional required bandwidth and improved speech quality. In this work we propose, as data protection technique, a piggybacking based forward error correction (FEC) method to be jointly used with the perceptual selection algorithm presented in Section II. The use of FEC techniques for speech data protection in error-prone channels avoids heavy latencies caused by Automatic Repeat-reQuest (ARQ) based strategies [28], thus resulting more suitable for real-time voice communications.

A secondary output of the algorithm described and discussed in the previous section is a perceptual based classification of the segments sent in each data packet. When a segment is selected for the transmission its energy E_i reflects its importance within the packet (bigger energy values indicate a bigger perceptual importance) and its position in a list of the most important perceptual segments. The main idea behind the proposed data protection strategy consists in protecting only the most perceptual important segments by means of a double transmission. More in detail, when the segments of the i_{th} packet are sent, these are stored and then duplicated according to their perceptual importance into the $(i+1)_{th}$ packet until reaching the maximum data payload size at the MAC layer. The $(i+1)_{th}$ packet is in charge of piggybacking the redundant information to be used to recover losses. The proposed strategy has the advantages to be easy to be implemented and fully compliant with the IEEE802.15.4 standard, although a delay of one packet is necessary in decoding a voice packet. The transport of both regular and duplicated segments through IEEE802.15.4 networks does not need any change at the MAC layer, although two fields must be introduced as first bytes of the MAC data payload to guarantee a correct reconstruction of the received duplicate segments. In Fig. 4 an IEEE802.15.4 packet for sending voice data according to the perceptual based selection and protection strategies is depicted. The two fields C_{mask} and P_{mask} are 2 bytes long and contain a mask indicating the positions of both current and redundant speech segments into an 80 bytes full speech packet. The N_{seg} value is supposed to be defined a priori, or when a call is established, thus avoiding to include it in every packet. Due to the overhead of the two introduced fields, the available MAC data payload

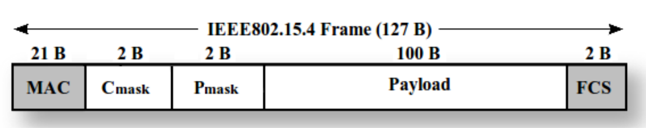

Fig. 4. IEEE802.15.4 data packet supporting the protection strategy.

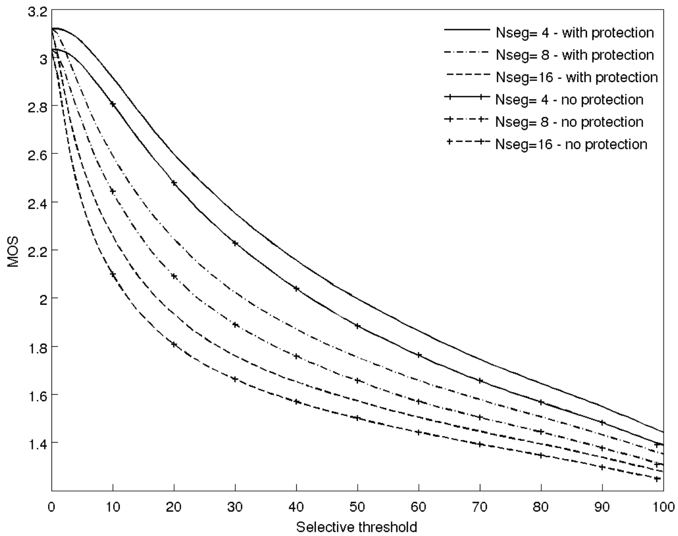

Fig. 5. MOS as a function of the selective threshold.

size is equal to 100 bytes.

B. Performance evaluation in a multi-hop scenario

The performance of the perceptual based data protection strategy has been evaluated by means of a simulative study in a multi-hop transmission scenario. The voice flow is supposed to be sent from the sender to the receiver in three hops in which packet loss can occur independently. In all the performed simulations real packet loss traces have been adopted. The used loss traces have been collected in a set of indoor experiments performed at the Institute of Communication, Information and Perception Technologies in Pisa. In the data collection setup several devices mounting an IEEE802.15.4 compliant transceiver have been installed in a corridor at a distance of 10 m while adopting a transmission power equal to 0 dBm. The selected environment is affected by signal attenuation, due to the movement of people, and interference, due to installed IEEE802.11 based access point [29], which results in a packet loss rate (PLR) equal to 1.66%, 1.77% and 2.24% for the traces adopted in the performed simulations. In case of voice communications without perceptual selection and protection the final PLR (with the three-hops communication) results equal to 5.50% with a MOS value equal to 3.03 against a value of 4.38 when no packet losses occur.

In Fig. 5 the performance in MOS values of the proposed protection strategy has been reported compared to a transmission in which no protection is applied. The results have been plotted as a function of the selective threshold S_{tr} for N_{seg} equal to 4, 8 and 16. The case N_{seg} equal to 2 has not been considered because of the lack of residual payload space for applying the protection strategy. Due to the proposed perceptual selection strategy, in which at least one segment of a packet is sent, the number of the transmitted packets is the same in any condition, as well as the experienced PLR, thus guaranteeing a fair comparison among protection and non-protection techniques for any selective threshold and number of segments. A first analysis of the results in Fig. 5 shows how the use of the proposed protection technique guarantees to reach higher MOS values with respect to a non-protected transmission for any value of the selective threshold at any imposed N_{seg} value. The MOS gain ranges from values close to zero to values close to 0.17 with the bell-shaped behavior plotted in Fig. 6. The MOS curves shapes reflect the same behavior of the bandwidth occupancy difference between protected and non-protected transmission reported in Fig. 7. Bigger MOS gains are reached with a bigger additional bandwidth. Regarding the bell-shaped behavior in the Fig. 6 and 7 this is due to two main effects: the bigger amount of available MAC data payload for protecting segments when S_{tr} increases and the low number of segments to be sent after a certain threshold value.

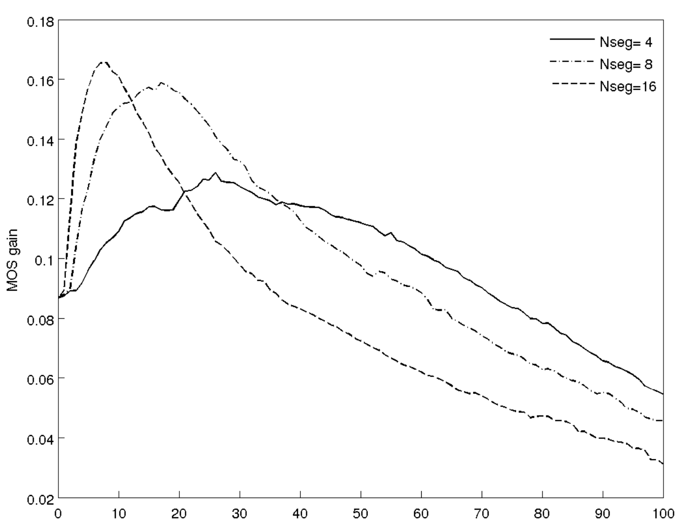

Fig. 6. MOS gain as a function of the selective threshold.

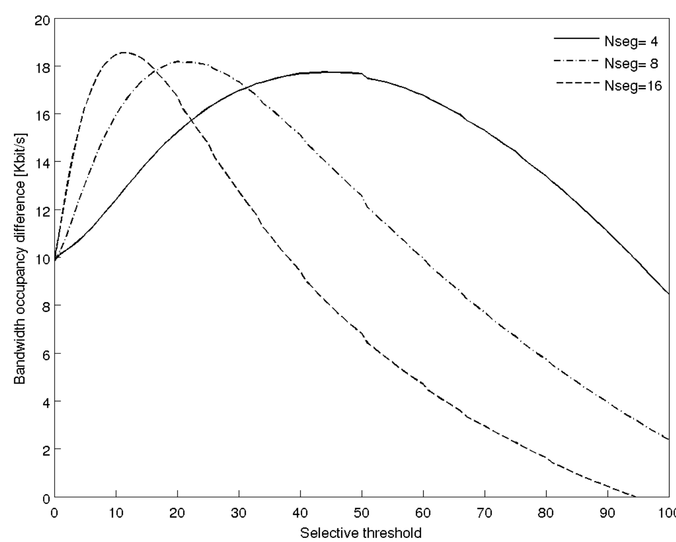

Fig. 7. Bandwidth difference as a function of the selective threshold.

The proposed protection strategy shows its benefits in particular for low values of the threshold when for selecting the segments to be sent high values of MOS are imposed to compensate the effect of a heavy PLR which dramatically reduces the end-to-end speech quality. In case low values of MOS can be accepted at the receiver side the use of the protection technique can be avoided while increasing the amount of the transmitted data (lower values of the S_{tr} threshold) in the perceptual selection. To be noted, finally, that the effect of both perceptual selection and protection techniques results in bigger values of MOS with respect to a transmission without selection and protection (3.12 against 3.03 in the performed simulations) still requiring lower values of bandwidth at MAC layer.

IV. Conclusions

In this paper the problem of enabling multiple voice communications in IEEE802.15.4 networks for emergency management support is analyzed. Starting from the consideration that the importance of a speech data packet, or part of it, to be sent through the network is function of its perceptual impact on the received speech stream, a perceptual based voice data selection technique is developed and presented. The aim of the developed selection technique is that of reducing the speech flow bandwidth, thus enabling multiple voice communications in WSN compliant with the IEEE802.15.4 standard, while preserving as much as possible the end-to-end speech quality. Performance evaluation results show as the developed perceptual selection algorithm can reach a 30.8% reduction in bandwidth occupancy with respect to a full rate voice communication still maintaining an end-to-end speech quality between good and fair according to the MOS scale.

The proposed speech selection algorithm is used in the paper as key part of a proposed forward error correction data protection technique able to recover speech data losses in real voice communications through IEEE802.15.4 networks. Simulation results based on real loss traces collected in an indoor environment show as the proposed protection technique provides its main benefits when high speech quality values have to be reached, reaching a MOS gain up to 0.17, while still saving useful bandwidth with respect to a full rate voice communication.

References

[1] J.P. Conti, "The Internet of Things," *Communications Engineer*, vol. 4, no. 6, pp. 20–25, January 2006.

[2] "Internet of Things in 2020: Roadmap for the Future," http://www.smart-systems-integration.org/public/internet-of-things, September 2008.

[3] L. Chen, Z. Chen, and S. Tu, "A realtime dynamic traffic control system based on wireless sensor network," in *Proceedings of International Conference Workshops on Parallel Processing*, Washington, USA, June 2005, pp. 258–264.

[4] W. Chen, L. Chen, Z. Chen, and S. Tu, "WITS: A wireless sensor network for intelligent transportation system," in *Proceedings of International Multi-Symposiums on Computer and Computational Sciences*, Hangzhou, China, June 2006, pp. 635–641.

[5] P. Pagano, M. Petracca, A. Alessandrelli, and C. Nastasi, "Enabling Technologies and Reference Architecture for a EU-Wide Distributed Intelligent Transport System," in *Proceedings of 8th ITS European Congress*, Lyon, France, June 2011.

[6] A. Alessandrelli, A. Azzarà, M. Petracca, C. Nastasi, and P. Pagano, "ScanTraffic: Smart Camera Network for Traffic Information Collection," in *Proceedings of European Conference on Wireless Sensor Networks*, Trento, Italy, February 2012, pp. 196–211.

[7] D. Snoonian, "Smart Buildings," *IEEE Spectrum*, vol. 40, no. 8, pp. 18–23, August 2003.

[8] K. Framling, I. Oliver, J. Honkola, and J. Nyman, "Smart spaces for ubiquitously smart buildings," in *Proceedings of International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, Sliema, Malta, October 2009, pp. 295–300.

[9] B. Bach, D. Wilhelmer, and P. Palensky, "Smart buildings, smart cities and governing innovation in the new millennium," in *Proceedings of IEEE International Conference on Industrial Informatic*, Osaka, Japan, July 2010, pp. 8–14.

[10] D. Han and J. Lim, "Smart home energy management system using IEEE 802.15.4 and zigbee," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 3, pp. 1403–1410, August 2010.

[11] E. Asimakopoulou and N. Bessis, "Buildings and crowds: Forming smart cities for more effective disaster management," in *Proceedings of International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, Seoul, Korea, July 2011, pp. 229–234.

[12] D. Efstathiou, S. Ioannou, S. Nikoletseas, and D. Patroumpa, "Emergency building evacuation guided by a wireless sensor network," in *Proceedings of International symposium on Mobility management and wireless access*, Miami, USA, November 2011, pp. 203–206.

[13] R. Mangharam, A. Rowe, R. Rajkumar, and R. Suzuki, "Voice over sensor networks," in *Proceedings of IEEE Real-Time Systems Symposium*, Rio de Janeiro, Brazil, December 2006, pp. 291–302.

[14] L. Li, G. Xing, L. Sun, and Y. Liu, "QVS: Quality-aware Voice Streaming for Wireless Sensor Networks," in *Proceedings of IEEE International Conference on Distributed Computing*, Montreal, Canada, June 2009, pp. 450–457.

[15] M. Petracca, G. Litovsky, A. Rinotti, M. Tacca, J.C. De Martin, and A. Fumagalli, "Perceptual based Voice Multi-Hop Transmission over Wireless Sensor Networks," in *Proceedings of IEEE Symposium on Computers and Communications*, Sousse, Tunisia, July 2009, pp. 19–24.

[16] IEEE Computer Society, "Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPAN)," *The Institute of Electrical and Electronics Engineers, Inc.*, October 2003.

[17] B. Latrè, P. De Mil, I. Moerman, B. Dhoedt, and P. Demeester, "Throughput and Delay Analysis of Unslotted IEEE 802.15.4," *Journal of Networks*, vol. 1, no. 1, pp. 20–28, May 2006.

[18] T.R. Park, T.H. Kim, J.Y. Choi, S. Choi, and W.H. Kwon, "Throughput and energy consumption analysis of IEEE 802.15.4 slotted CSMA\CA," *Electronics Letters*, vol. 41, no. 18, pp. 1017–1019, September 2005.

[19] ITU-T Recommendation G.711, "Pulse code modulation (PCM) of voice frequencies," November 1988.

[20] J.D. Gibson, *Multimedia Communications*, chapter 1, pp. 1–12, Academic Press, San Diego, USA, November 2000.

[21] ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality," August 1996.

[22] ITU-T Recommendation P.862, "Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone net-works and speech codecs," February 2001.

[23] J.C. De Martin, "Source-driven packet marking for speech transmission over differentiated-services networks," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Salt Lake City, USA, August 2002, pp. 753–756.

[24] A. Bayya and M. Vis, "Objective Measures for Speech Quality Assessment in Wireless Communications," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, Atlanta, USA, May 1996, pp. 495–498.

[25] NTT Advanced Technology Corporation (NTT-AT), "Multi-Lingual Speech Database for Telephonometry," 1994.

[26] ITU-T Recommendation P.862.1, "Mapping function for transforming P.862 raw result scores to MOS-LQO," November 2003.

[27] L. Ding and R.A. Goubran, "Assessment of effects of packet loss on speech quality in VoIP," in *Proceedings of IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications*, Ottawa, Canada, September 2003, pp. 49–54.

[28] I.F. Akyildiz, T. Melodia, and K.R. Chowdury, "A Survey on Wireless

Multimedia Sensor Networks," *Computer Networks*, vol. 51, no. 4, pp. 921–960, March 2007.

[29] H. Khaleel, C. Pastrone, F. Penna, M.A. Spirito, and R. Garello, "Impact of Wi-Fi traffic on the IEEE 802.15.4 channels occupation in indoor environments," in *Proceedings of International Conference on Electromagnetics in Advanced Applications*, Torino, Italy, September 2009, pp. 1042–1045.