Video multicast using unequal error protection with Luby Transform codes

Shakeel Ahmad, Raouf Hamzaoui, Marwan Al-Akaidi
Faculty of Technology, De Montfort University, Leicester, UK.
Email:{sahmad,rhamzaoui,mma}@dmu.ac.uk

Abstract— **In a recent work, we proposed a method for unequal error protection with Luby Transform codes and showed that it achieves lower bit error rates than a state of the art technique when the information symbols are partitioned into two protection levels (most important and least important). In this paper, we apply our previous work to the problem of video multicast with heterogeneous receivers. We provide simulations for the scalable video coding (SVC) extension of the H.264/AVC standard and show that our unequal error protection method provides significantly better objective video quality results than two state of the art techniques in applications where a high video quality is desired.**

Keywords: Channel coding, fountain codes, unequal error protection, video multicast, performance analysis.

I. INTRODUCTION

Fountain codes [1], [2], [3] are a new class of erasure codes that have two advantages over traditional erasure codes such as Reed-Solomon codes. First, they allow faster encoding and decoding, making them more suitable for real-time applications. Second, they are more flexible in the sense that the channel code rate does not have to be fixed in advance as an infinite number of encoded symbols can be built on the fly.

Luby Transform (LT) codes [2] were the first class of practical Fountain codes. Shokrollahi [3] introduced another class of practical Fountain codes called Raptor codes by concatenating a fixed-rate channel code with an LT code. Raptor codes have been adopted as enhanced application layer forward error correction (FEC) by Multimedia Broadcast/Multicast System (MBMS) of the 3rd Generation Partnership Project (3GPP), IP datacast (IPDC) of Digital Video Broadcasting - Handheld (DVB-H), as well as Digital Video Broadcasting Project's (DVB) global IPTV standard.

Recently, some works [4], [5], [6], [7] have addressed the problem of designing Fountain codes with unequal error protection (UEP) properties. In UEP, information symbols are protected according to their importance. This usually allows a better overall system performance than equal error protection where all information symbols receive the same level of protection [8].

In [7], we provided a simple method to decrease the bit error rate (BER) of LT codes by virtually increasing the number of information symbols. Moreover, we exploited this idea to propose a new technique for UEP with LT codes and compared the BER performance of our scheme to that of [4]. However, our results were provided for general data without a specific application in mind. Moreover, our UEP scheme was not compared to another state of the art technique [5]. In this paper, we apply the work in [7] to the problem of video multicast with heterogeneous receivers. We provide experimental results for the scalable video coding (SVC) [9] extension of the H.264/MPEG-4 AVC video compression standard and show that our method provides significantly better peak signal to noise ratio (PSNR) results than the state of the art techniques of [4] and [5] in applications where a high video quality is desired.

The paper is organized as follows. Section II contains background material about LT codes. Section III describes the UEP techniques of [4], [5], and [7]. Section IV presents our simulation results for the video multicast scenario.

II. BACKGROUND

In this section, we explain the encoding and decoding process with LT codes.

A. Encoding

The LT encoder takes a set of k information symbols (bits or bytes, for example) and generates a potentially infinite sequence of encoded symbols of the same alphabet. Each encoded symbol is computed independently of the other encoded symbols. More precisely, given k information symbols $i_1, \dots, i_k$ and a suitable probability distribution $\Omega(x)$ on $\{1, \dots, k\}$, a sequence of encoded symbols $e_n,\ n \geq 1, \dots,$ is generated as follows. For each $n \geq 1$

1) Select randomly a degree $d_n \in \{1, \dots, k\}$ according to the distribution $\Omega(x)$.
2) Select uniformly at random d_n distinct information symbols and set e_n equal to their bitwise modulo 2 sum.

The relationship between the information symbols and encoded symbols can be described by a graph (see Fig. 1 for an example).

B. Decoding

When an encoded symbol is transmitted over an erasure channel, it is either received correctly or lost. The LT decoder tries to recover the original information symbols from the received encoded symbols. We assume that for each received encoded symbol, the decoder knows the indices of the information symbols it is connected to. This is possible, for example, by using a pseudo-random

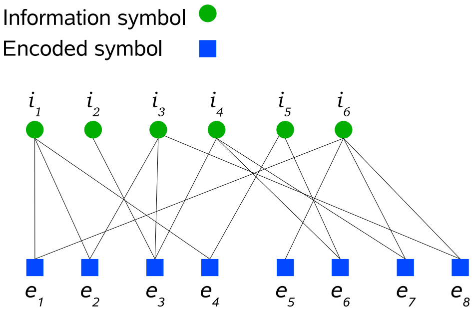

Fig. 1. Encoding graph of an LT code. Eight encoded symbols are generated from $k = 6$ information symbols. The degree of an encoded symbol is the number of information symbols that were used to generate it. For example, the degree of e_1 is equal to two.

generator with the same seed as the one used by the encoder.

The decoding process is as follows:

1) Find an encoded symbol e_m that is connected to only one information symbol i_j. If this is not possible, stop the decoding.
   a) Set $i_j = e_m$.
   b) Set $e_x = e_x \oplus i_j$ for all indices $x \neq m$ such that e_x is connected to i_j. Here $\oplus$ denotes the bitwise modulo 2 sum.
   c) Remove all edges connected to i_j.
2) Go to Step 1.

III. PREVIOUS UEP METHODS

In this section, we describe the three existing UEP techniques with LT codes.

A. Rahnavard, Vellambi, and Fekri's method [4]

Rahnavard, Vellambi, and Fekri [4] were the first to propose a method to provide UEP with LT codes. For simplicity, we describe their method when two levels of protection are used. Consider a source block having k information symbols. Partition these k information symbols into two sets S_1 and S_2 of size $|S_1| = \alpha k$ and $|S_2| = (1 - \alpha)k$, respectively, where $0 < \alpha < 1$. The set S_1 is called the set of most important bits (MIB) while the set S_2 is called the set of least important bits (LIB). Define probabilities p_1 and p_2 ($p_1 + p_2 = 1$) to select S_1 and S_2, respectively. Given a suitable probability distribution $\Omega(x)$ on $\{1, \ldots, k\}$, a sequence of encoded symbols e_n, $n \geq 1$ is generated as follows. For each n

1) Select randomly a degree $d_n \in \{1, \ldots, k\}$ according to the distribution $\Omega(x)$.
2) Select d_n distinct information symbols successively. To select a symbol, first select one of the two sets S_1 or S_2 (S_1 with probability p_1 and S_2 with probability p_2). Then choose randomly a symbol from the selected set.
3) Set e_n equal to the bitwise modulo 2 sum of the d_n selected information symbols. Figure 2 describes the process for two classes.

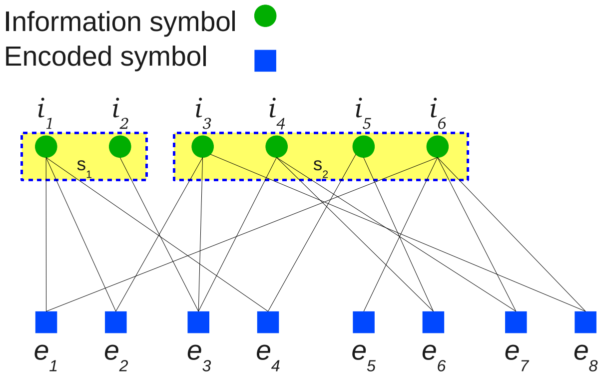

Fig. 2. UEP scheme proposed in [4]. Two levels of protection are used. The set of MIB contains two information symbols while the set of LIB contains four information symbols.

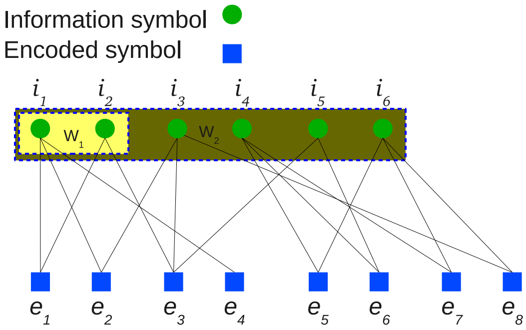

Fig. 3. UEP scheme proposed in [5]. Two levels of protection are used. The encoded symbols e_1 and e_4 are generated from the MIB class while the remaining encoded symbols are generated from the LIB class.

To ensure that the MIB symbols have lower BER than the LIB symbols, the probability of selecting an MIB symbol should be larger than the probability of selecting an LIB symbol [4], that is, $p_1 \frac{1}{|S_1|} > p_2 \frac{1}{|S_2|}$. To achieve this, one can set $p_1 = \frac{k_M |S_1|}{k}$ and $p_2 = \frac{k_L |S_2|}{k}$ for $0 < k_L < 1$ and $k_M = (1 - (1 - \alpha) k_L)/\alpha$. Here the parameter k_M gives the relative importance of the MIB symbols.

B. Method of Sejdinovic et al. [5]

A source block having k information symbols is partitioned into L sets $S_1, S_2, \cdots, S_L$ such that the first $|S_1|$ information symbols of the source block belong to the set of most important bits, the next $|S_2|$ information symbols belong to the set of next most important bits and so on. Then L windows $W_1, W_2, \cdots, W_L$ are defined such that W_i consists of the sets $S_1, \cdots, S_i$. Thus the size of the ith window is $|W_i| = \sum_{j=1}^{i} |S_j|$ and $|W_1| < |W_2| < \cdots < |W_L|$. To generate an encoded symbol, a window is selected according to distribution $\Gamma(x) = \sum_{i=1}^{L} \Gamma_i x^i$. Here Γ_i is the probability that window W_i is chosen. Then an LT code with a suitable degree distribution is applied. UEP is achieved by choosing appropriate values for $\Gamma_1, \cdots, \Gamma_L$. Fig. 3 describes the encoding process for two classes.

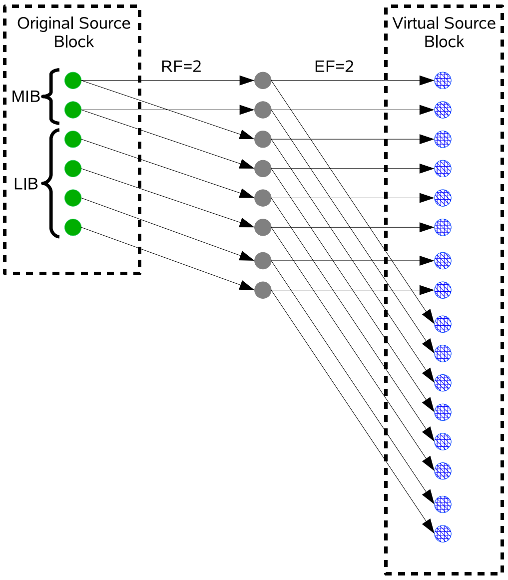

Fig. 4. Building a virtual source block with the UEP scheme proposed in [7]. Here $4k = 6$, $EF = 2$, and $RF = 2$.

C. Ahmad, Hamzaoui, and Al-Akaidi's method [7]

For simplicity and without loss of generality, we describe our UEP method on an example where there are two levels of protection (MIB and LIB symbols), $k = 6$ information symbols, and two MIB symbols. If we duplicate the two MIB symbols as in the first step of Fig. 4, the virtual size of the source block becomes 8, corresponding to a repeat factor $RF = 2$. Here RF is the number of times a block of MIB symbols is present in the virtual source block. We next extend the degree distribution of the LT code from $k = 6$ to $k = 8$. To generate an encoded symbol, we find its degree d using the new degree distribution and then select d information symbols from the 8 virtual symbols. If the index of a selected information symbol is larger than 2, we map its virtual index to the actual index by subtracting 2. This UEP technique can be combined with block duplication [7] by duplicating the virtual source block with an expanding factor EF. For example, for $RF = 2$ and $EF = 2$, the original source block consisting of two MIB symbols and four LIB symbols is transformed into a virtual block of size $EF(RF \times 2 + 4) = 16$ (Fig. 4).

IV. EXPERIMENTAL RESULTS

In this section, we compare our UEP technique to the UEP techniques of [4] and [5] for a video multicast scenario with heterogeneous receiver classes. In this scenario, a video server multicasts an SVC [9] encoded bitstream of $(1+t)k$ symbols to n receiver classes. The parameter t is called the transmission overhead.

As in [6], we used the Stefan video sequence which has a spacial resolution of 352×288 and a temporal resolution of 30 fps. The first group of pictures (GOP) of the sequence was encoded using the SVC reference software (JVSM) into one base layer (BL) and 14 enhancement layers (ELs). The GOP had a length of 16 frames. The resulting layer sizes, bit rates, and PSNR are summarized in Table I. A source block consisted of one GOP, and its size was about 190,000 bytes.

Decoded layers	Size	Bitrate	PSNR
BL	400	292.37	25.79
BL + 1 EL	700	510.65	27.25
BL + 2 EL	875	636.56	28.14
BL + 3 EL	1155	839.82	29.00
BL + 4 EL	1550	1127.10	29.51
BL + All ELs	3800	2764.55	40.28

TABLE I

SVC ENCODING OF THE FIRST GOP OF THE STEFAN VIDEO SEQUENCE (352X288,30 FPS) INTO ONE BASE LAYER (BL) AND 14 ENHANCEMENT LAYERS (EL). THE TABLE SHOWS THE NUMBER OF SYMBOLS, THE BITRATE IN KBPS, AND THE Y-PSNR IN dB.

As in [6], each symbol was equal to 50 bytes, giving $k = 3800$ symbols per source block. In each source block, the base layer (containing 400 symbols) was chosen as the MIB set, whereas the remaining symbols built the LIB set. The video server transmitted 250 source blocks. At the receiver side, we assumed that a layer can be used to enhance the video quality only if it was decoded fully, and all the layers before this layer were also decoded fully. In this way, the number of consecutively decoded information symbols, starting from the first information symbol, determined the number of decoded layers.

For the UEP scheme of [4], we followed the recommendation in [4] and used $k_M = 2$ and the fixed degree distribution of [3] for both the MIB and LIB sets. This degree distribution is given by

$$\begin{aligned}\Omega(x) =& 0.007969x + 0.493570x^2 + 0.166220x^3 + \\ & 0.072646x^4 + 0.082558x^5 + 0.056058x^8 + \\ & 0.037229x^9 + 0.055590x^{19} + 0.025023x^{64} + \\ & 0.003135x^{66}.\end{aligned} \tag{1}$$

For the UEP scheme of [5], we followed the recommendation in [6] and used $\Gamma_1 = 0.11$, the fixed degree distribution (1) for the LIB class, and the robust soliton distribution [2] with parameters $c = 0.03$ and $\delta = 0.5$ for the MIB class.

For our UEP scheme, we used the robust soliton distribution with parameters $c = 0.01$ and $\delta = 0.5$.

Fig. 5 shows the PSNR performance of our UEP scheme as a function of the symbol loss rate for various settings of RF and EF. Here the transmission overhead t was equal to one, and the video was transmitted to one receiver. The results show that increasing RF from 2 to 3 increases the likelihood of decoding the BL successfully, but decreases the likelihood of successfully decoding the ELs, leading to a decrease of the overall PSNR. Increasing EF increases the average degree and leads to duplicate selection of the information symbols, which decreases the

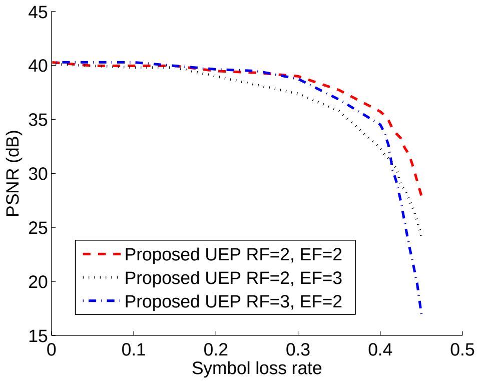

Fig. 5. PSNR as a function of the symbol loss rate for the transmission of the Stefan sequence with the proposed UEP scheme. The performance of the scheme is shown for different settings of EF and RF.

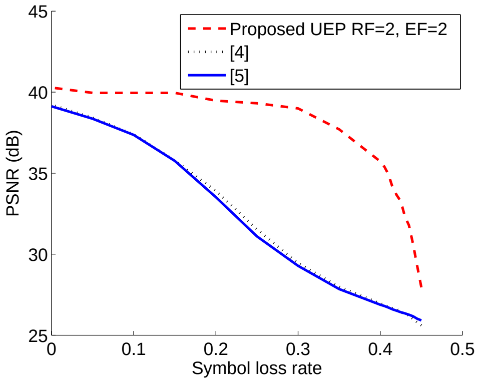

Fig. 6. PSNR as a function of the symbol loss rate for the proposed UEP scheme, the UEP scheme of [4], and the UEP scheme of [5].

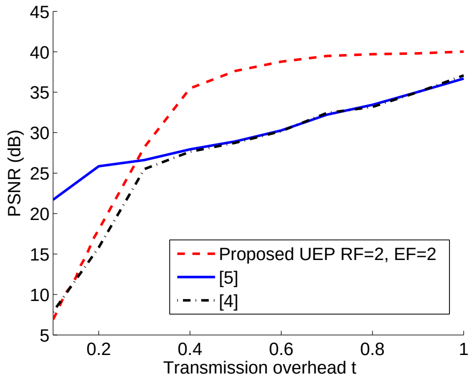

Fig. 7. Average PSNR of $n = 5$ receiver classes as a function of the transmission overhead.

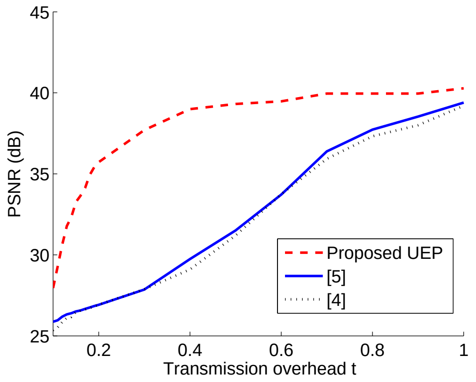

Fig. 8. PSNR of receiver class with symbol loss rate 0 as a function of the transmission overhead.

performance. Based on these results, we set $RF = 2$ and $EF = 2$ for the following experiments.

Fig. 6 shows the PSNR perceived by a receiver class as a function of the symbol loss rate for our UEP scheme, the UEP scheme of [4], and the UEP scheme of [5]. Our scheme was more robust to increasing symbol loss rate.

Fig. 7 shows the average PSNR performance for a multicast transmission to $n = 5$ receiver classes with symbol loss rates 0, 0.05, 0.1, 0.15, and 0.2. Our UEP scheme always outperformed the UEP scheme of [4]. It also outperformed the UEP scheme of [5] over the range of transmission overheads that are required to provide full decoding of the BL. In particular, our UEP scheme achieved an average video quality of 36 dB with 60 % less transmission bandwidth.

Fig. 8, 9, and 10 show the individual PSNR results for receiver classes with symbol loss rates 0, 0.1, and 0.2, respectively. Compared to our UEP scheme, the scheme of [5] achieved an acceptable video quality (25 dB) for the receiver with the worst channel conditions at a lower overhead (Fig. 10). However, with the scheme of [5], the maximum PSNR for this receiver did not exceed 32.5 dB even at $t = 1$. In contrast, our scheme reached this PSNR for a transmission overhead t=0.48, exceeding 35 dB at $t = 1$.

V. CONCLUSION

We used our LT-based UEP scheme [7] for video transmission over a lossy channel. Simulations for the SVC extension of the H.264 standard showed that our scheme can provide up to 7 dB improvement in PSNR over the UEP schemes of [4], [5] for unicast transmission. For multicast transmission to a set of receivers with different channel conditions, simulations showed that our UEP scheme has a better average PSNR performance when the transmission overhead is large and a worse performance when the overhead is low. This makes our UEP scheme more appropriate in applications where high video quality is desired. In this paper, the values of RF and EF were optimized by simulations. Future work could be the development of faster optimization techniques based on analytical models. Other future work could extend the proposed framework to the situation where the information symbols are partitioned into more than two sets.

VI. ACKNOWLEDGMENT

This work was supported in part by the DFG Research Training Group GK-1042. We thank Vladimir Stanković for providing a draft of paper [6].

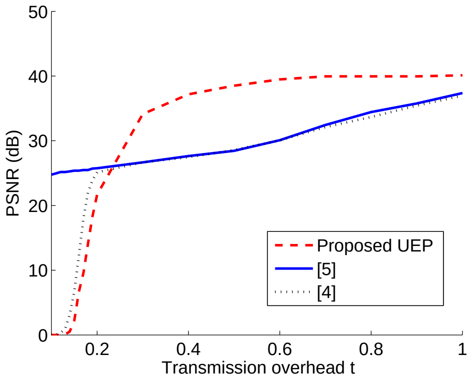

Fig. 9. PSNR of receiver class with symbol loss rate 0.1 as a function of the transmission overhead.

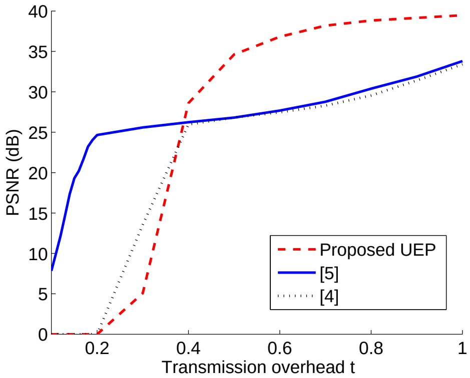

Fig. 10. PSNR of receiver class with symbol loss rate 0.2 as a function of the transmission overhead.

REFERENCES

[1] J.W. Byers, M. Luby, and M. Mitzenmacher, “A digital fountain approach to asynchronous reliable multicast”, IEEE Journal on Selected Areas in Communications, vol. 20, pp. 1528–1540, Oct. 2002.

[2] M. Luby, “LT-Codes”, Proc. 43rd Annual IEEE Symposium on Foundations of Computer Science, pp. 271–280, 2002.

[3] A. Shokrollahi, “Raptor Codes”, IEEE Trans. Inf. Theory, Vol. 52, No. 6, June 2006, pp. 2551–2567.

[4] N. Rahnavard, B. N. Vellambi, and F. Fekri, “Rateless codes with unequal error protection property”, IEEE Trans. Inf. Theory, Vol. 53, No. 53, pp. 1521–1532, April 2007.

[5] D. Sejdinovic, D. Vukobratovic, A. Doufexi, V. Senk, and R. Piechocki, “Expanding window fountain codes for unequal error protection”, Proc. 41st Asilomar Conf., Pacific Grove, pp. 1020–1024, 2007.

[6] D. Vukobratovic, V. Stanković, D. Sejdinovic, L. Stanković, and Z. Xiong, “Scalable video multicast using expanding window fountain codes”, IEEE Transactions on Multimedia, Special Issue on Quality-driven Cross-layer Design for Multimedia Communications, vol. 11, August 2009.

[7] S. Ahmad, R. Hamzaoui, and M. Al-Akaidi, “Unequal error protection using LT codes and block duplication”, Proc. MESM 08, Middle Eastern Multiconference on Simulation and Modelling, Amman, Aug. 2008.

[8] R. Hamzaoui, V. Stanković, and Z. Xiong, “Optimized error protection of scalable image bitstreams, IEEE Signal Proc. Mag., vol. 22, pp. 91-107, Nov. 2005.

[9] H. Schwarz, D. Marpe, and T. Wiegand, “Overview of the Scalable Video Coding Extension of the H.264/AVC Standard”, IEEE Trans. Circuits and Systems for Video Tech. vol. 17, pp. 11031120, September 2007.

BIOGRAPHIES

Shakeel Ahmad received the B.Sc. degree in Electrical Engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2000, the M.Sc. degree in Information and Communication Systems from Hamburg University of Technology (TUHH), Germany, in 2005, and the Dr.-Ing degree from the University of Konstanz, Germany, in 2008. He is currently working as a Research Fellow with the Faculty of Technology at De Montfort University, Leicester, U.K. His research interests include multimedia communication, error-resilient video streaming and channel coding.

Raouf Hamzaoui received the Maîtrise de mathématiques from the Faculty of Sciences of Tunis, Tunis, Tunisia, in 1986, the M.Sc. degree in mathematics from the University of Montreal, Montreal, QC, Canada, in 1993, the Dr. Rer. Nat. degree from the Faculty of Applied Sciences, University of Freiburg, Freiburg, Germany, in 1997, and the Habilitation degree in computer science from the University of Konstanz, Konstanz, Germany, in 2004. He is currently a Professor of Media Technology in the Faculty of Technology at De Montfort University, Leicester, U.K. His research interests include image and video compression, multimedia communications, error control systems, and algorithms.

Marwan M. Al-Akaidi (PhD CEng CIP FIEE FBCS FIAP SMIEEE) is a Professor and research group leader of the Mobile & Wireless Multimedia Comms & SP group. He obtained his first degree (82) in Electronic Engineering (UK) and his MSc & PhD in Digital Communications in 84 & 88 respectively from Loughborough University, UK. He led the school of Engineering & Technology at De Montfort University from Jan. 2002 to Sept. 2008, and currently is appointed as the Director of External relations. He enjoys strong links with UK & Europe industries including Siemens, O2, BT, Orange, Nokia, BMW, Chronos and many others. He is interested in all aspects of Mobile & Wireless Communications & signal processing (including Speech & video) and he is the author of the ”Fractal Speech Processing” book (Cambridge press). He is the SPC chair & conference coordinator of the IEEE UKRI. He is also the chair of the IET 3G Technology & beyond 3G technology conferences as well as the WiMAX08 London, WiMAX09 London conferences and the chair & founder of the MESM conferences since 99. He is also editor-in-chief of AWMC & editorial member of many national & International Journals.