EXTREME UNEQUAL ERROR PROTECTION AND PERFECTLY RELIABLE ENCODING

Riccardo Bernardini and Roberto Rinaldo

DIEGM–University of Udine
Via delle Scienze 208, 33100 Udine, Italy
{riccardo.bernardini,rinaldo}@uniud.it

ABSTRACT

In this paper we define the new concept of a *perfectly reliable encoder* (PRE) for erasure channels as an encoder whose performance is optimal for every value of the loss probability. As a first step toward the design of such an encoder, we propose an idealized *extreme unequal error protection* encoding scheme. We derive models to optimize the encoder performance. We show by experiments that the proposed scheme approaches the theoretical behavior, and quantify the gap with respect to a scheme specifically designed for a given loss probability.

Index Terms— multimedia encoding, reliable encoding, unequal error protection, erasure channel

1. INTRODUCTION

Multimedia encoding for transmission over unreliable channels has stimulated a lot of research activity in the last few years. Basically, the idea behind every scheme for reliable coding is to add some redundancy (e.g., via multiple description, erasure codes, and so on) so that the receiver can decode the content with good quality even in the presence of errors (typically, packet losses). The trade-off between the added redundancy and the base quality of the content (i.e., the quality perceived when no packet is lost) is handled differently by different reliable encoding schemes.

In this paper we approach the reliable encoding problem from a different point of view, asking ourselves if a *Perfectly Reliable Encoder (PRE)* exists. To understand what we mean with "perfectly reliable encoder" consider the following setup: some multimedia content is encoded at an overall rate (that is, including erasure correction data) of R_{tot} bit/sample. The content is sent over a broadcast channel (e.g., wireless) and received by many receivers. Every receiver experiences different channel conditions represented, for example, by different loss probabilities P_ℓ. We will say that the encoder is *perfectly reliable* if every client experiences the best quality achievable for its channel conditions.

Note that in the described setup the content is encoded *only once* and encoding parameters (e.g., how much redundancy is added) *cannot* be adapted to each channel condition. Therefore, although it is easy to construct an encoder that gives the optimal performance for a fixed loss probability P_ℓ, in this case we want an encoder that it is (ideally) optimal *simultaneously* for all possible loss probabilities. It is worth observing that a PRE would be useful not only for transmission over erasure channels. For example, a PRE would make rate adaptation very easy. In order to adapt the content to the available bandwidth it would suffice to throw away data at random; the same procedure could also be used, for example, to do congestion control when sending data over UDP.

In this paper, we propose to investigate if a PRE can be actually designed, and we analyze the performance of an idealized Unequal Error Protection scheme that could seem a natural choice toward this objective. To the best of our knowledge, the analysis and the transmission setup considered in this paper are not fully analyzed in the literature, despite the great interest for applications.

In particular, we consider an abstract scheme, which we call *Extreme Unequal Error Protection (EUEP)*, where the signal is encoded with an embedded coder whose output bitstream is partitioned into many pieces, each one protected with an erasure code. The length of the pieces and the redundancy used for each piece are optimized so that the distance between the performance of the encoder and the theoretical best curve is minimized. The proposed scheme is clearly ideal, so that its performance is an upper bound to the performance of every practical scheme based on Unequal Error Protection (UEP). It is shown that the best choice gives rise to a curve that it is an offset version of the ideal one, a nice behavior, albeit with a performance price to pay.

1.1. Prior Work

The literature on reliable coding is wide and variegate. In the last years many approaches have been proposed including Multiple Description Coding (MDC) using linear spaces approaches (e.g., subsampling, frame expansions, correlating transforms) [1] – [6], special quantizers [7], or Forward Error Correction (FEC) codes [8] – [17], and others. In some case (e.g., [18]) FEC-based reliable encoding is used in cross-layer structures. Reliable encoding has been proposed for a variety of applications: video/image transmission over lossy networks is the most common one, but also video transmission over wireless links [18], transmission over peer-to-peer networks [19] and others. All the cited works but [15] share the characteristic that they work with a fixed channel, that is, they suppose that the statistical description of the channel is given.

Differences with Existing Literature. Since the approach analyzed here is FEC-based, we will consider only existing literature proposing FEC-based solutions. The works proposing UEP solutions can be roughly partitioned into two classes: papers that propose a specific coder for some specific application [9, 10, 18, 19] and more theoretical papers that analyze in more generality the problem of optimizing a UEP scheme, without any specific reference to the type of encoded signal [8, 11, 12, 13, 14, 15, 16, 17]. Because of the theoretical nature of this paper, we will consider only the second group.

As already said, most of the existing works in the second group consider a fixed channel whose statistical description, described in terms of the probability p_n of receiving n packets, is known. Almost all those works use numerical procedures to find the best UEP for a given channel; the exceptions are [17] that proposes a closed-form solution, but only for a special type of R-D curve, and [13] that proposes a sub-optimal, but efficient, progressive algorithm. Our work differs from those works because we *do not* try to optimize the UEP for a given channel, but we consider the problem of minimizing the maximum difference between the distorsion experienced by a client whose memoryless erasure channel has a loss probability equal to P_ℓ and the optimal distorsion that the client could experience with a specifically designed encoder. In other words, we do not consider a specific channel, but we try to optimize for *all* the possible P_ℓ.

From this point of view, [15] is the paper whose setup most closely resemble ours. In [15] the authors consider the problem of sending the same content, but at different resolutions, to many clients, with the objective of minimizing the maximum difference between the quality perceived by each client and the corresponding optimal quality. They propose a numerical procedure specifically tailored to this problem based on norm-infinity minimization, as it is done in this paper. However, we do not consider the problem of serving different clients that require different resolutions and, moreover, we do not minimize over a set of channels, but over all the possible P_ℓ.

2. PROBLEM STATEMENT

Suppose a transmitter can encode a multimedia content with a budget rate (that includes both content and protection data) equal to R_{tot} bit/sample. Suppose the encoded content is broadcast and received by a specific client through an erasure memoryless binary channel with loss probability P_ℓ. It is well known that, whatever the encoding mechanism, the best SNR that this client can achieve is $\text{SNR}(P_{\text{rec}}R_{\text{tot}})$ where

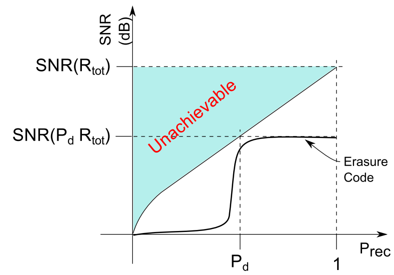

Fig. 1. Bound on the quality achievable with rate at the source equal to R_{tot} and packet reception probability equal to P_{rec}.

$P_{\text{rec}} := 1 - P_\ell$ is the probability of receiving a bit and $\text{SNR}(R)$ denotes the rate-vs-SNR function associated with the multimedia content, that is, $\text{SNR}(R)$ is the maximum SNR (in dB) theoretically achievable with a rate equal to R bit/sample (in absence of losses). In other words, the best quality that the receiver can experience is the quality associated with the "actual rate" $P_{\text{rec}}R_{\text{tot}}$ seen by the receiver. Fig. 1 shows the graph of the ideal P_{rec}-vs-SNR curve, which approaches a straight line as P_{rec}, and hence the effective rate, increases. Note that the region above the curve is not achievable.

As well known, every point of the ideal curve is asymptotically achievable by using a suitable erasure code. Fig. 1 also shows a (qualitative) example of the P_{rec}-vs-SNR curve achievable with an erasure code. The curve exhibits the typical "cliff effect": the code allows for a full recovery of the data up to a given code design probability P_d. For loss probabilities larger than P_d, the probability of recovering the data gets smaller and the performance decays sharply. Note that at $P_{\text{rec}} = 1$ the SNR is lower than $\text{SNR}(R_{\text{tot}})$ because some of the rate budget is used by the code.

A PRE is an encoder whose P_{rec}-vs-SNR curve can be made as close as desired to the ideal curve shown in Fig. 1, possibly at the cost of an increased complexity. Since UEP schemes are potential candidates for this type of problem, we will analyze in the following an idealized form of UEP coding.

3. EXTREME UNEQUAL ERROR PROTECTION

The scheme that we analyze in this paper is an abstract scheme that can be thought as an extreme version of an UEP scheme. Given its ideal nature, the performance that we will obtain will be an upper bound to the one of any practical UEP scheme. In this paper we will suppose that we can use a *perfectly embedded encoder*.

Hypothesis 1. *A* perfectly embedded encoder *is available, that is, an encoder whose output bit-stream is such that, by decoding only the first H bits of the bit-stream obtained by encoding N samples, one obtains a Signal-to-Noise ratio equal to* $\text{SNR}(H/N)$.

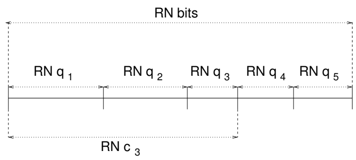

Fig. 2. Partitioning of the bit-stream

3.1. Encoding

Let N be the number of samples of the information content and let R_{tot} be the maximum number of bit per sample that we can transmit. The idea is to encode the content with the perfectly embedded encoder at a rate of $R < R_{\text{tot}}$ bit/samples, obtaining a string of RN bits, and using the remaining $(R_{\text{tot}} - R)N$ bits for protection. The protection bits should be optimally distributed to approach the limit curve of Fig. 1. Let $\rho := R_{\text{tot}}/R$ denote the overall redundancy.

More into detail, the bit-string produced by the encoder is partitioned into L consecutive blocks where the i-th block $(i = 1, \ldots, L)$ has size $q_i RN$. Obviously, it must be $\sum_{i=1}^{L} q_i = 1$. It will be convenient to denote with c_n the "cumulative sums" of the first n values q_i, that is, $c_n = \sum_{i=1}^{n} q_i$ (with, of course, $c_0 = 0$). Note that $c_i RN$ is the length of the bit-string corresponding to the first i blocks. Fig. 2 shows a graphical representation of the notation.

We protect the i-th block with a redundancy equal to $r_i > 1$, that is, we add $(r_i - 1)q_i RN$, redundancy bits by using an erasure code (e.g., a Reed-Solomon code). Therefore, we use $r_i q_i RN$ bits for the i-th block and we will be able to recover the i-th block if and only if we receive at least $q_i RN$ bits out of $r_i q_i RN$. The constraint on the total bit budget requires that $\sum_{i=1}^{L} r_i q_i = R_{\text{tot}}/R = \rho$.

3.2. Decoding

The decoder operates in a standard way: if the first k blocks have been recovered, but not the $k+1$-th one, the decoder recovers the signal using the first k blocks and discards the others (even if some block after the $k+1$-th is correctly received). Note that in this case the decoder decodes using $c_k R_{\text{tot}}/\rho$ bit/sample, with an SNR equal to $\text{SNR}(c_k R_{\text{tot}}/\rho)$, in accordance with the hypothesis that the encoder is perfectly embedded.

Our objective is to determine the values of q_i and r_i such that the performance of this scheme is as close as possible (in a sense to be defined) to the best theoretical performance. If necessary, we will allow both N (the content size) and L (the number of blocks) to go to infinity.

3.3. Decoder performance

For the sake of notational convenience, denote with $\mathscr{E}_k$ the event corresponding to the correct reception of the first k blocks, but not the $k+1$-th one ($\mathscr{E}_L$ corresponds, clearly, to the reception of all the blocks). Observe that events $\mathscr{E}_k$, $k = 0, \ldots, L$, are a partition of the probability space, so that the average distortion at the receiver $\mathbb{E}[D]$ can be obtained by conditioning over $\{\mathscr{E}_k\}_{k=0}^{L}$. If $\mathscr{P}_k(P_{\text{rec}})$ denotes the probability of $\mathscr{E}_k$ when the probability of reception is P_{rec}, one can write

$$\mathbb{E}[D] = \sum_{k=0}^{L} E[D|\mathscr{E}_k]\mathscr{P}_k(P_{\text{rec}}) = \sum_{k=0}^{L} D_k \mathscr{P}_k(P_{\text{rec}}) \qquad (1)$$

where D_k is the distortion obtained when $\mathscr{E}_k$ happens. Since $\mathscr{P}_k(P_{\text{rec}})$ is a complex function of P_{rec}, (1) is difficult to analyze. However, it is possible to approximate (1) with a simpler expression that gets close to (1) when N goes to infinity.

The idea is that if N is large, because of the properties of the erasure code, we expect to be able to recover segment k if $P_{\text{rec}} > p_k := 1/r_k$ and not recovering it if $P_{\text{rec}} < p_k$. In other words, we expect that if $p_k < P_{\text{rec}} < p_{k+1}$ the event $\mathscr{E}_{k+1}$ will happen with probability close to 1. As a consequence, the probability $\mathscr{P}_k(P_{\text{rec}})$ can be approximated with

$$\mathscr{P}_k(P_{\text{rec}}) \approx \chi_k(P_{\text{rec}}) \qquad (2)$$

where $\chi_k(x) = 1$ if $p_k \leq x < p_{k+1}$. By using (2) in (1) we obtain a step-wise approximation to $\mathbb{E}[D]$

$$\mathbb{E}[D] \approx \sum_{k=0}^{L} D_k \chi_k(P_{\text{rec}}) \qquad (3)$$

Intuitively, we expect that approximation (3) will "get better" as N increases. In order to make this statement more precise, we define a measure of "closeness" mutated, in spirit, from filter design.

Definition 1. *Let $I \in \mathbb{R}$ be an interval, let $S = \{s_1, \ldots, s_A\} \subset I$ be a set of points of I, let $\varepsilon > 0$ and let $S_\varepsilon := \cup_{k=1}^{A}[s_k - \varepsilon, s_k + \varepsilon]$. We will say that $f : I \to \mathbb{R}$ and $g : I \to \mathbb{R}$ are (S, ε)-*close*, denoted as $f \underset{S,\varepsilon}{\sim} g$, if*

$$\sup_{x \notin S_\varepsilon} |f(x) - g(x)| \leq \varepsilon \qquad (4)$$

Informally, f and g are (S, ε)-close if they do not differ more than ε everywhere, save on set S_ε that plays a role similar to transition bands in filter design. See Fig. 3.

Property 1. *For every choice of $\{q_1, \ldots, q_L\}$, $S := \{p_1, \ldots, p_L\}$ and ε it is possible to find N large enough such that*

$$D(P_{rec}) = \sum_{k=0}^{L} D_k \mathscr{P}_k(P_{rec}) \underset{S,\varepsilon}{\sim} \sum_{k=0}^{L} D_k \chi_k(P_{rec}) \qquad (5)$$

The proof of Property 1 is not difficult, but it is long and it is not given here. Since intervals $[p_k, p_{k+1})$ are disjoint, we can write the "operative" SNR, SNR_{op} as

$$\text{SNR}_{\text{op}}(p) = \sum_{k=0}^{L} \text{SNR}(c_k R_{\text{tot}}/\rho + 1)\chi_k(p) \qquad (6)$$

that is, SNR_{op} is a step-wise function too. See Fig. 4

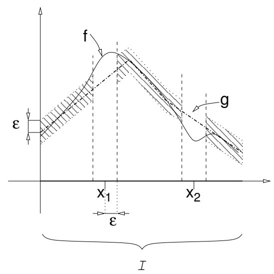

Fig. 3. Example of two functions f and g that are (S,ε)-close with $S = \{x_1, x_2\}$.

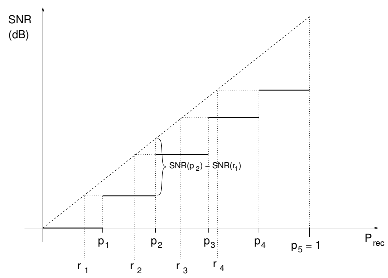

Fig. 4. Step-wise approximation of the P_{rec}-vs-SNR_{op} curve and illustration of the distance between the SNR_{op} and SNR_{th}.

3.4. Optimal Encoder Design

In order to find the "best" EUEP encoder we need a metric to measure the distance between the operative SNR function (6) and the best theoretical signal-to-noise ratio. Since SNR_{op} in (6) is a function of p, it is convenient to rescale the best theoretical SNR in order to make it function of p too. Therefore, we define $\text{SNR}_{\text{th}} : [0,1] \to \mathbb{R}$, $\text{SNR}_{\text{th}}(x) := \text{SNR}(xR_{\text{tot}})$. Our objective function is

$$\Delta_{\text{SNR}} := \|\text{SNR}_{\text{th}} - \text{SNR}_{\text{op}}\|_\infty \tag{7}$$

Since SNR_{op} is step-wise, the maximum error is achieved on the step borders, so that (see also Fig. 4)

$$\Delta_{\text{SNR}} = \max_{k=1,\dots,L} \text{SNR}(p_{k+1}R_{\text{tot}}) - \text{SNR}(c_k R_{\text{tot}}/\rho), \tag{8}$$

where we set $p_{L+1} = 1$. At large bit-rates $\text{SNR}(x) \approx ax + b$, for some $a, b \in \mathbb{R}$ and (8) becomes

$$\Delta_{\text{SNR}} \approx aR_{\text{tot}} \max_{k=0,\dots,L} p_{k+1} - \frac{c_k}{\rho}.$$

Since $a > 0$, we deduce that minimizing (8) is approximately equivalent to minimize

$$\Delta_{\text{p}} := \max_{k=0,\dots,L} p_{k+1} - \frac{c_k}{\rho} \tag{9}$$

that depends only on q_i and p_i.

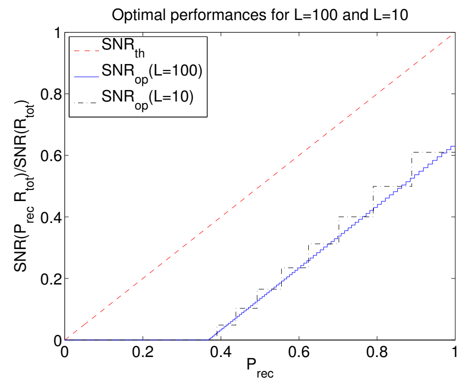

Fig. 5. Plot P_{rec}-vs-SNR of the best solution for $L = 100$ and $L = 10$ compared with the best theoretical performance.

3.4.1. Numerical optimization

We carried out the numerical optimization of (9) using the `fminimax` function of ®Matlab for several values of L under the constraints $0 \le p_1 < \cdots < p_L < p_{L+1} = 1$, $q_i \ge 0$, $\sum_{i=1}^{L} q_i = 1$. Fig. 5 shows the results obtained for[1] $L = 100$ and $L = 10$, with a comparison with the theoretical limit. In Fig. 5 the vertical axis has been normalized by $\text{SNR}_{\text{th}}(R_{\text{tot}})$ so that the range of the vertical axis is $[0,1]$ and the ideal curve is (in the large bit-rate approximation) a 45-degree line. As it can be seen, the performance curve of the best EUEP is parallel to the theoretical limit, with an horizontal offset approximately equal to 0.37, almost insensitive to the value of L. The performance for $L = 10$ are slightly worse than for $L = 100$, since the maximum distance from the theoretical curve is larger. This number quantifies the price one has to pay in order to protect simultaneously against different P_ℓ. In particular, for a certain P_{rec}, the user will experience the same quality as in system specifically designed to protect errors corresponding to $P'_{\text{rec}} = P_{\text{rec}} - 0.37$.

4. CONCLUSIONS

We defined the new concept of *perfectly reliable encoder* (PRE) in the context of erasure channels and proposed an *extreme unequal error protection* (EUEP) encoding scheme as a possible approach toward the design of a PRE. We derived models that are used to optimize the performance of EUEP. We found that the best EUEP has a P_{rec}-vs-SNR curve parallel to the optimal curve. This shows that the performance of the best EUEP is qualitatively similar to the theoretical limit, with a relatively high gap with respect to a system designed for a particular P_{rec}. Future research will explore the possibility to use other approaches to the design of a PRE, in order to possibly reduce this gap.

[1] Similar results are obtained for L ranging from 10 up to thousands.

5. REFERENCES

[1] P. L. Dragotti, S. D. Servetto, and M. Vetterli, "Optimal Filter Banks for Multiple Description Coding: Analysis and Synthesis," *IEEE Transactions on Information Theory*, vol. 48, pp. 2036–2052, July 2002.

[2] V. K. Goyal and J. Kovačević, "Generalized multiple description coding with correlating transforms," *IEEE Transactions on Information Theory*, vol. 47, Sept. 2001.

[3] Riccardo Bernardini, Marco Durigon, Roberto Rinaldo, Luca Celetto, and Andrea Vitali, "Polyphase spatial subsampling multiple description coding of video streams with h264," in *Proceedings of ICIP 2004*, Singapore, Oct. 2004, pp. 3213–3216.

[4] P. L. Dragotti, Jelena Kovačević, and V. K. Goyal, "Quantized oversampled filter banks with erasures," in *Proc. IEEE Data Compression Conference*, Snowbird, UT, Mar. 2001, pp. 173–182.

[5] V. K. Goyal, Jelena Kovacevic, and J. A. Kelner, "Quantized frame expansions with erasures," *Applied and Computational Harmonic Analysis*, vol. 10, no. 3, pp. 203–233, May 2001.

[6] R. Bernardini, M. Durigon, R. Rinaldo, A. Vitali, and P. Zontone, "Frame-based multiple-description video coding with extended orthogonal filter banks," *EURASIP Journal on Applied Signal Processing*, vol. 2006, pp. Article ID 53623, 17 pages, 2006.

[7] S. D. Servetto, V. A. Vaishampayan, and N. J. A. Sloane, "Multiple description lattice vector quantization," in *Proc. IEEE Data Compression Conference*, Snowbird, UT, Mar. 1999, IEEE, pp. 13–22.

[8] Jason Quinlan, Ahmed Zahran, and Cormac Sreenan, "SDC: Scalable description coding for adaptive streaming media," in *Packet Video Workshop (PV), 2012 19th International*. IEEE, 2012, p. 59–64, IEEE, n/a.

[9] M.R. Ardestani and A.A. Beheshti Shirazi, "Multiple description coding of scalable video stream based on unequal loss protection and optimal rate allocation," *Scientia Iranica*, pp. 782–294, June 2011, doi:10.1016/j.scient.2011.06.010.

[10] Tammam Tillo, Enrico Baccaglini, and Gabriella Olmo, "Unequal protection of video data according to slice relevance," *IEEE Transactions on Image Processing*, vol. 20, no. 6, pp. 1572–1582, June 2011, doi:10.1109/TIP.2010.2095865.

[11] Shakeel Ahmad, Raouf Hamzaoui, and Marwan M. Al-Akaidi, "Unequal error protection using fountaing codes with applications to video communication," *IEEE Transactions on Multimedia*, vol. 13, no. 1, pp. 92–101, Feb. 2011, doi:10.1109/TMM.2010.2093511.

[12] Sorina Dumitrescu, Geoffrey Rivers, and Shahram Shirani, "Unequal erasure protection technique for scalable multistreams," *IEEE Transactions on Image Processing*, vol. 19, no. 2, pp. 422–434, Feb. 2010, doi:10.1109/TIP.2009.2032943.

[13] Majid Ardestani, Alireza Aminlou, and Ali Ashgar Beheshti Shirazi, "Unequal loss-protected multiple description coding of scalable source streams using a progressive approach," in *Proc. ICIP 2009*, 2009, pp. 901–904.

[14] Sorina Dumitrescu, Xiaolin Wu, and Zhe Wang, "Globally optimal uneven error-protected packetization of scalable code streams," *IEEE Transactions on Multimedia*, vol. 6, no. 2, pp. 230–239, Apr. 2004, doi:10.1109/TMM.2003.822793.

[15] Vladimir Stanković, Raouf Hamzaoui, and Zixiang Xiong, "Robuts layered multiple description coding of scalable media data for multicast," *IEEE Signal Processing Letters*, vol. 12, no. 2, pp. 154–157, Feb. 2005, doi:10.1109/LSP.2004.840895.

[16] Alexander Mohr, Eve Riskin, and Richard Ladner, "unequal loss protection: graceful degradataion of image quality over packet erasure channels through forward error correction," *IEEE Journal on Selected Areas in Communication*, vol. 18, no. 6, pp. 819–828, June 2000.

[17] R. Puri and K. Ramchandran, "Multiple description source coding using forward error correction codes," in *Signals, Systems, and Computers, 1999. Conference Record of the Thirty-Third Asilomar Conference on*, Aug. 2002, vol. 1, pp. 342–346, doi:10.1109/acssc.1999.832349.

[18] Wei-Ho Chung, Sunil Kumar, Seethal Paluri, Santosh Nagaraj, Annamalai Annamalai Jr., and John Matyajs, "A cross-layer unequal error protection scheme for prioritized H.264 video using RCPC codes and hierarchical QAM," *J. Inf. Process Syst.*, vol. 9, no. 1, pp. 53–66, Mar. 2013.

[19] Yuanyuan Xu, Ce Zhu, Wenjun Zeng, and Xue Jun Li, "Multiple description coded video streaming in peer-to-peer networks," *Signal Processing: Image Communication*, pp. 412–429, Feb. 2012, doi:10.1016/j.image.2012.02.005.