

RESOURCE MANAGEMENT FOR SCALABLY ENCODED INFORMATION: THE CASE OF IMAGE TRANSMISSION OVER WIRELESS NETWORKS

Virgilio Rodriguez

ECE Department, Polytechnic University
5 Metrotech Center, Brooklyn, NY 11201
E-mail:vr@ieee.org

ABSTRACT

Scalable encoded information, as in the JPEG 2000 standard, results in files which can be truncated at an arbitrary point and decoded. This work introduces a tractable, yet flexible model appropriate for resource management involving scalably encoded information. At its core is a function yielding a measure of “quality” of the decoded information as a function of the number of bits chosen for decoding. It is assumed that all that is known about this function is that its graph yields an “S-curve”. An energy-efficient policy for the transmission over a wireless network of scalably-encoded images is sought. Two variables are jointly optimized: transmission power, and the number of bits of each file to be decoded (“coding rate”). The single-user case is fully analyzed, and a closed-form solution given, which can be clearly identified, graphically. The analysis indicates that both variables can be “decoupled”, and their optimal values found independently of each other.

1. INTRODUCTION

At the foundations of the JPEG 2000 image compression standard there are ideas found in the embedded zero-tree wavelet coding (EZW) algorithm introduced by [6], a technique which produces a fully “embedded” bit stream[7]. An embedded bit stream is “scalable”, in the sense that it can be truncated at an arbitrary point, and decoded. If bits are decoded as they are received, at any instant the “quality” of the decoded information is the best available for the number of bits received up to that moment. Thus, an image compression ratio can be varied simply by truncating the coded bit stream. Similar ideas can be applied to video coding. In fact, fine-granular scalability (FGS) is at the core of the MPEG-4 video-compression standard.

Scalable coding can be fruitfully exploited in many practical applications, including: (i) image database browsing (ii) progressive image transmission (where the consumer

can examine the improving decoded image as bits are received, and can abort the transfer when the image quality reaches a satisfactory level), and (iii) multimedia web serving (a single file can serve a variety of consumer requirements and capabilities, as well as congestion/channel conditions).

These files introduce interesting resource management issues, because their special structure can be exploited to allocate scarce resources efficiently. Such analysis necessitates a relatively simple model combining the properties of analytical tractability, with flexibility to accommodate a wide variety of situations. This work proposes such model.

In the situation under study, a terminal with a limited supply of energy and a long sequence of scalably encoded images to transfer over a wireless link seeks to manage its energy efficiently. At the center of this inquiry is a function yielding the “quality” of the resulting information (image) in terms of the fraction of the encoded file which is chosen for decoding. It is postulated that all that is known about this function is that its graph is an S-curve, as introduced in [4] and discussed further in [5] (see fig. 1).

There are practical reasons why the S shape is chosen. An arbitrary S-curve starts out convex and smoothly transitions to concave. But the inflexion (transition) point is arbitrarily placed. Therefore, this curve in fact contains as special cases a (“mostly”) concave curve (inflexion point is “very close” to the origin) and a (“mostly”) convex curve (inflexion point is “very far” from the origin). Thus, by assuming an S shape for the function giving the “quality” of the image recovered from the truncated file in terms of the number of decoded bits, this work allows not only the S-shape proper, but also the concave and the convex shape. These three shapes should accommodate most, if not all situations of interest.

Additionally, as discussed in fundamental psychology texts (see, for instance, [2, Chapter 7]), the S-curve naturally arises in psychophysical experiments involving human perception. In these experiments, a graph is made in which the horizontal axis denotes the “intensity” of a stimulus applied to a subject. The vertical axis denotes the probability

Supported in part by NYSTAR through the Wireless Internet Center for Advanced Technology (<http://wicat.poly.edu>) at Polytechnic Univ.

that the subject correctly identifies or detects the presence of the stimulus. These graphs have often the shape of fig. 1.

The peak signal-to-noise ratio (PSNR) is the image quality metrics most commonly found in the literature. This is a simple to calculate index, which can be sensible and useful in many situations. However, as an indicator of image quality as perceived by a human observer, the PSNR is at best a very crude measure. Dansereau and Kinsner [1] argues this further, while proposing a metrics specifically aimed at progressive image transmission: the Renyi dimension spectrum. But this measure is much too complicated for resource management studies.

This analysis also depends critically on a function giving the probability of success of the transmission of a data packet in terms of a signal to interference measure at the receiver. This “frame-success” function (FSF) is determined by the physical layer of the system. It can be safely assumed that for any physical layer, any such function has an S-shaped graph. Thus, two different S-curves are at the core of this analysis.

The single-terminal case is fully analyzed, and the foundation is laid for a multi-terminal analysis. The problem is set up as a joint optimization in which two key variables are jointly optimized: transmission power, and the number of bits of each file to be decoded. A closed-form solution is given.

The scientific literature contains various works involving power allocation and the transmission of scalably encoded information. The most relevant may be [3], which considers files which have been “layer coded” (a form of discrete scalable coding) and seeks a power allocation policy across the various layers, minimizing the overall end-to-end distortion. However, previous works seeking a joint power, and coding rate selection in order to maximize an image quality metrics appear unavailable.

2. CONCEPTUAL FRAMEWORK

2.1. Quality as a function of the number of decoded bits

At the center of this inquiry is a function yielding the quality of the decoded image as a function of the number of bits in the fraction of the encoded file which is decoded. Image quality is a subjective matter. Nevertheless, certain basic assumption can be made about the properties of a function giving quality as a function of received bits. About this function, it is postulated that:

1) Its domain is the interval $[0, M]$, where M is the length in bits of the entire encoded file.

2) Its range is the interval $[0, 1]$. This is just a normalization. A 1 denotes the best possible quality of the decoded image (say the quality of the original).

3) It must be strictly increasing (more decoded bits yield a better image quality, by design)

4) Its graph is S-shaped, as in figure (1). In practical terms, sigmoidness further implies that: (a) If the number of decoded bits is sufficiently large, the quality of the decoded image will be sufficiently close to “perfect”. (b) After a sufficiently large number of bits have been decoded, the marginal contribution to image quality of an additional bit becomes “very small” and is decreasing. (c) If the number of decoded bits is sufficiently small, the quality of the decoded image will be sufficiently close to zero. (d) Bits at the beginning of the encoded file contribute to the perceived “quality” of the image at an increasing rate (“initial convexity”). One plausible interpretation is that even a highly distorted image may provide enough information to identify its “meaning” (what is it? a bird?, a person’s face?, etc.). This essential semantic information is provided by the bits at the beginning of the encoded file (“base layer”).

References [4, 5] discuss the technical characterization of a generic S-shaped function. A fixed function could work for different images, in particular if the images are sufficiently “similar” (e.g., each corresponds to a passport picture of a respective adult).

2.2. A Generalized frame-success function

The frame-success function (FSF) yields the probability that a data packet is received successfully as a function of the received signal to interference ratio. This function is determined by physical attributes of the system, including the modulation technique, the forward error detection scheme, the nature of the channel, and properties of the receiver. It is assumed that all that is known about the FSF, f_s , is that its graph exhibits a sigmoidal shape as in figure (1). More specifically, it is assumed that the function defined by $f(x) = f_s(x) - f_s(0)$ obeys the properties of the generalized sigmoidal function introduced in [4] and discussed further in [5].

3. SINGLE-USER ANALYSIS

3.1. Problem statement

The problem faced by a single transmitter in a wireless (in particular CDMA) network can be formulated as follows.

It is taken as given a (1) certain amount of energy, E , available for transmission, (2) fixed transmission rate of R bits per second, (3) long sequence of files each of length M , each divided into blocks of bits (packets/frames) of length $L \ll M$ and each corresponding to equally important *similar* images encoded scalably, (4) function g as defined in section 2.1 giving the quality of an image obtained by decoding a truncated encoded file as a function of the numbers of bits decoded, (5) certain level of interference (noise), (6) function f_s as described in section 2.2 giving the probability that a data frame is received successfully as a function

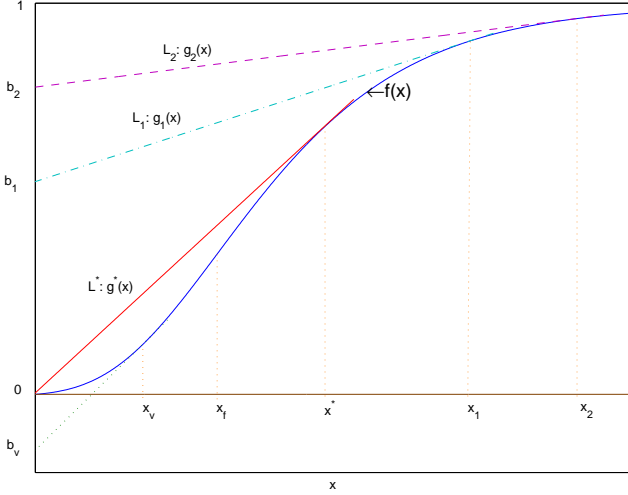


Fig. 1. An “S-curve” and some of its tangents

of the signal to interference ratio at the receiver.

The transmitter wants to choose optimally (i) the number of successfully received bits at which point a given file can be considered successful, so that the transmission of the next file is started (that is, the “optimal” level of image quality at which point it is considered “good enough”), and (ii) the transmission power. The objective is to maximize the weighted number of images transferred by the time the available energy runs out, where the weight is the quality of each image. This criterion can also be stated as maximizing the “total quality” transferred.

Because packets have been assumed much smaller than a file, the fact that the number of bits in a frame and file is an integer is ignored. Because the images are similar enough (e.g. each image corresponds to a (respective) human face), the same function works for all images.

3.2. Objective function

Suppose that at a certain instant of time, $y < M$ bits of the current file have been received. Then, $g(y) \leq 1$ gives the quality of the image that would result if the file containing the received bits is decoded.

Let $Q = P \cdot h$ be the power at the receiver when a certain data packet is transmitted with power P ; and let I be the interference (noise) power. Then, $f_s(GQ/I)$ is the probability that said packet is correctly received (G is a CDMA constant, the spreading/processing gain).

Assuming that, once a packet is received in error, re-transmission is ideal, then the total number of times a given packet needs to be (re)-transmitted is a geometric random variable, whose probability distribution is of the form $\pi(1 - \pi)^{k-1}$, with $\pi = f_s(GQ/I)$. The expected value of this

random variable is $1/\pi$, interpreted as the average number of times the same packet needs to be transmitted to ensure correct reception.

The average amount of energy that needs to be spent in order to achieve the successful reception of an image of quality $g(y)$ when transmission power is set to P can be obtained as follows. Each packet requires an amount of energy equal to the product of 3 factors: the power P , the length in time of a packet (given the transmission rate R), and the average number of times the same packet needs to be transmitted to ensure correct reception. Each L -bit packet lasts L/R secs. Therefore, the average amount of energy required by a packet is $P \cdot (L/R) \cdot (1/\pi)$. Since y bits of data contain y/L packets, the average amount of energy necessitated by the successful reception of an image of quality $g(y)$ is given by

$$\frac{PL}{\pi R} \cdot \frac{y}{L} = \frac{Py}{\pi R} \quad (1)$$

To obtain the average number of images of quality $g(y)$ which can be successfully transmitted with an energy budget E , we divide E by the preceding expression (eq. (1)), and obtain:

$$\frac{\pi RE}{Py} \equiv RE \frac{f_s(GhP/I)}{Py} \quad (2)$$

To obtain the total received “image quality”, the preceding expression needs to be multiplied by $g(y)$, the quality of each image. Therefore, the user wants to maximize

$$RE \frac{f_s(GhP/I)g(y)}{Py} = RE \frac{f_s(GhP/I)}{P} \frac{g(y)}{y} \quad (3)$$

For technical reasons discussed in [5], $f_s(x)$ is replaced with $f(x) = f_s(x) - f_s(0)$. Then we can re-write equation () as

$$\frac{REGh}{I} \frac{f(GhP/I)}{GhP/I} \frac{g(y)}{y} \propto \frac{f(x)}{x} \frac{g(y)}{y} \quad (4)$$

with $x := GhP/I$.

3.3. Optimization Model and Solution

In view of the preceding analysis, the objective of the single user can be summarized as

$$\max \frac{f(x)}{x} \frac{g(y)}{y} \quad (5)$$

$$\text{s.t. } 0 \leq y \leq M \quad (6)$$

$$0 \leq x \leq \bar{x} \quad (7)$$

where $\bar{x} := Gh\bar{P}/I$ with \bar{P} the largest available transmission power.

Notice that the ratios in the objective function (5) are mutually independent; i.e., one does not influence or constraint the other. Therefore, the ratios $f(x)/x$ and $g(y)/y$ can be maximized independently, and the maximum of the

product of the ratios can be obtained as the product of the individual maxima. This problem can be easily solved by invoking the results provided by [4], and discussed in [5]. Reference [4] finds the maximum of the ratio $f(x)/x$ s.t. $0 \leq x \leq \bar{x}$ where f all that is known about f is that its graph is S-shaped. The maximizer is the lesser of \bar{x} and x^* . x^* is the abscissa of the unique point where the graph of f is tangent to a ray emanating from the origin (See figure (1)).

From the preceding paragraph, the maximum of $f(x)/x$ s.t. $0 \leq x \leq \bar{x}$ is obtained at $x^{**} = \min \{x^*, \bar{x}\}$. Likewise, the maximum of $g(y)/y$ s.t. $0 \leq y \leq M$ is obtained at $y^{**} = \min \{y^*, M\}$. The single-user problem is solved.

4. DISCUSSION

The problem faced by an energy-limited terminal with a long list of scalably encoded similar images to transfer over a wireless link has been solved. A tractable model, based on two “S-curves”, has been discussed. A closed-form solution is given in terms of a point which can be easily identified in the graph of the pertinent S-curve. The analysis leads to the maximization, over an appropriate region, of the product $Rf(\gamma)/P \times g(y)/y$, where γ is the received SIR, P is the transmission power, R the data transmission rate, f is the “frame success” function, y is the chosen number of decoded bits, and g is the “quality” function. $Rf(\gamma)/P$ has the unit bits/Joule, well known in the power control literature (see [5]), while quality/bit is the unit of $g(y)/y$. Hence, the maximized product is an intuitively appealing index in quality/Joule.

Although the problem is set up as a joint optimization of power and coding rate, the development indicates that both variables can be “decoupled”. In retrospect, this seems reasonable. The files are transmitted in small segments (data packets) which are assumed much smaller than the files, constant, and independent of y , the number of bits chosen for decoding. Power is needed to increase the probability that a data packet is received successfully. But the physical layer treats each packet in the same way, irrespective of the file to which it belongs, or its position within its file. Thus, the point, y , at which a given file is truncated to start the transmission of the next file has no effect on the probability of success of the intervening packets. Future research could investigate making packet length a variable dependent on y , a shorter packet length for a smaller y .

The S-curve practically contains as special cases a strictly convex and a strictly concave curve. However, it is shown in [5] that, if f (resp. g) were strictly concave, the ratio $f(x)/x$ (resp. $g(y)/y$) would be maximized at zero. In this

case, the power level, $\propto x$, (resp. the “truncation point”, y) should be set as small as possible. Likewise, if f (resp. g) were strictly convex, then the power level, (resp. the “truncation point”) should be set as large as possible.

This analysis can be extended to include many terminals sharing a CDMA channel. In this case, each terminal’s “noise” must include the interference caused by others. The problem can be set up as a “game” in which each terminal seeks to maximize its quality/Joule index. In this formulation, the key question is the existence and characterization of a “Nash equilibrium” (NE); i.e., a feasible allocation (of power and file size here) to each terminal, such that no terminal would be better off by *unilaterally* changing its allocation. Both of the ratios ($f(x)/x$ and $g(y)/y$) making up the quality/Joule index are quasi-concave[4]. It is well-known that a game in which “pay-off” functions are quasi-concave, and each player’s “strategy space” (power and file size here) is closed and bounded does have a NE. Game theory has been fruitfully applied to the transmission of conventional data over a wireless channel, in [5] and other works.

5. REFERENCES

- [1] Dansereau, R. and W. Kinsner, “Psychovisual correlations with multifractal measures for wavelet and wavelet packet progressive image transmission”, Can. Conf. on Elec. Comp. Eng., Vol. 1, pp. 435 -9, 2000
- [2] Gray, P., “Psychology”, New York:Worth Pub., 2002.
- [3] Khansari, M. and M. Vetterli, “Layered transmission of signals over power-constrained wireless channels”, IEEE ICIP, Vol. 3, pp. 380 -83, Oct 1995
- [4] Rodriguez, V., “Maximizing the Ratio of a Generalized Sigmoid to its Argument”, WICAT Tech. Rep. 02-010, Polytechnic Univ., 2002 <http://wicat.poly.edu/reports>
- [5] Rodriguez, V., “Robust Modeling and Analysis for Wireless Data Resource Management”, IEEE WCNC, 2003
- [6] Shapiro, J.M., “Embedded image coding using zerotrees of wavelet coefficients”, IEEE Trans. on Signal Proces., Vol. 41, No. 12, pp. 3445 -62, Dec 1993
- [7] Usevitch, B.E., “A tutorial on modern lossy wavelet image compression: foundations of JPEG 2000”, IEEE Signal Proc. Mag. , Vol. 18, No. 5, pp. 22 -35, Sep 2001