MP3 AND STREAMING AUDIO

Section 7.2 explained how an audible sound or a human voice ranging between 20 Hz and 20 KHz can be converted into digital bits and eventually into packets for networking. A signal is sampled, quantized, and encoded, a process called PCM. A variety of methods of compressing such encoded products at the output of the PCM are available. However, Huffman compression of the processed signal may not be sufficient for transmission over IP networks.

The MPEG-1 layer 3 (MP3) technology compresses audio for networking and producing CD-quality sound. The sampling part of PCM is performed at a rate of 44.1 KHz to cover the maximum of 20 KHz of audible signals. Using the commonly used 16-bit encoding for each sample, the maximum total bits required for audio is 16 x 44.1 = 700 kilobits and 1.4 megabits for two channels if the sound is processed in a stereo fashion. For example a 60-minute CD (3,600 seconds) requires about 1.4 x 3,600 = 5,040 megabits, or 630 megabytes. This amount may be acceptable for recording on a CD but is considered extremely large for networking, and thus a carefully designed compression technique is needed.

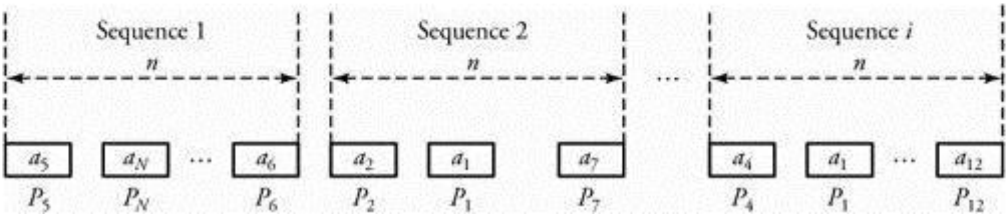
MP3 combines the advantages of MPEG with "three" layers of audio compressions. MP3 removes from a piece of sound all portions that an average ear may not be able to hear, such as weak background sounds. On any audio streaming, MP3 specifies what humans are not able to hear, removes those components, and digitizes the remaining. By filtering some part of an audio signal, the quality of compressed MP3 is obviously degraded to lower than the original one. Nonetheless, the compression achievement of this technology is remarkable.

Limits of Compression with Loss:

Hartely, Nyquist, and Shannon are the founders of information theory, which has resulted in the mathematical modeling of information sources. Consider a communication system in which a source signal is processed to produce sequences of n words, as shown in Figure 7.10. These sequences of digital bits at the output of the information source can be compressed in the source encoder unit to save the transmission link bandwidth. An information source can be modeled by a random process $X_n = (X_1, ..., X_n)$, where X_i is a random variable taking on values from a set of values as $\{a_1, ..., a_N\}$, called alphabet. We use this model in our analysis to show the information process in high-speed networks.

Figure 7.10. A model of data sequences

[View full size image]



Basics of Information Theory

The challenge of data compression is to find the output that conveys the most information. Consider a single source with random variable X, choosing values in $\{a_1, ... a_N\}$.

If a_i is the most likely output and a_j is the least likely output, clearly, a_j conveys the most information and a_i conveys the least information. This observation can be rephrased as an important conclusion: The measure of information for an output is a decreasing and continuous function of the probability of source output. To formulate this statement, let P_{k1} and P_{k2} be the probabilities of an information source's outputs a_{k1} and a_{k2}, respectively. Let $I(P_{k1})$ and $I(P_{k2})$ be the information content of a_{k1} and a_{k2}, respectively. The following four facts apply.

1. As discussed, $I(P_k)$ depends on P_k.
2. $I(P_k)$ = a continuous function of P_k.
3. $I(P_k)$ = a decreasing function of P_k.
4. $P_k = P_{k1}.P_{k2}$ (probability of two outputs happen in the same time).
5. $I(P_k) = I(P_{k1}) + I(P_{k2})$ (sum of two pieces of information).

These facts lead to an important conclusion that can relate the probability of a certain data to its information content:

Equation 17.16

$$I(P_k) = -\log_2 P_k = \log_2\left(\frac{1}{P_k}\right).$$

The log function has a base 2, an indication of incorporating the binary concept of digital data.

Source : http://elearningatria.files.wordpress.com/2013/10/cse-vi-computer-networks-ii-10cs64-notes.pdf