Foundations of Computer Security

Lecture 34: Fundamental Theorems

Dr. Bill Young
Department of Computer Sciences
University of Texas at Austin

Source Entropy

The entropy of a language is a measure of the most efficient possible encoding of the language.

The *entropy of a message source* is the amount of information content that the source can produce in a given period. This is measured in bits per second.

Any channel can transmit an arbitrary amount of information, *given enough time and unbounded buffering capacity*. But can a given channel transmit the information *in real time*?

Channel Capacity

The *capacity* of a channel is the number of bits that can be sent per second over the channel. This is a property of the communication medium.

Fundamental Theorem of the Noiseless Channel. (Shannon): If a language has entropy h (bits per symbol) and a channel can transmit C bits per second, then it is possible to encode the signal is such a way as to transmit at an average rate of $(C/h) - \epsilon$ symbols per second, where ϵ can be made arbitrarily small. It is impossible to transmit at an average rate greater than C/h.

What the Theorem Means

Suppose a channel can handle 100 bits / second and your language has entropy 5 (bit per symbol).

Given a perfect encoding and a noiseless channel, you'd expect to be able to transmit 20 symbols / second though the channel, on average. Right?

But you may not have a perfect encoding. Doesn't matter. You can always find a better encoding to get within ϵ of that limit.

Noisy Channels

If the channel is noisy, the capacity is reduced by the noise. But the following is true:

Fundamental Theorem of a Noisy Channel (Shannon): Let a discrete channel have a capacity C and a discrete source an entropy h (bits per second). If $h < C$ there exists a coding system such that the output of the source can be transmitted over the channel with an arbitrarily small frequency of errors.

If the channel (with noise factored in) can physically handle the message traffic, then it is possible to transmit with arbitrarily small error rate.

What this Theorem Means

The upshot of this is that a message can be transmitted reliably over even a very noisy channel by increasing the redundancy of the coding scheme.

For example, covert channels in the system cannot be dismissed with the argument that they are noisy and hence useless. You can always get the message through by finding a more redundant encoding.

Lessons

- Entropy provides a bound on coding efficiency.
- But Shannon's theorems show that it is always possible to approach that limit arbitrarily closely.

Next lecture: Entropy of English