Category Overview

8/14/07

Communication

Moving data reliably from one place to another

The Internet has become the world's largest distributed communication, retrieval, and computing system. Its basic functions, sending and securing messages, are probably as old as humanity. What's new is its global reach and virtually instantaneous delivery of messages.

Many methods of encoding messages have been invented for reliable communicating when voice cannot carry the distance or when privacy is needed. Native Americans used smoke signals to communicate over many miles, seafarers used signal semaphores ship-to-ship, and military commands developed ciphers to hide their communications from their enemies.

Technologies that facilitate and extend communication have always been winners. Perhaps the most touted example is the printing press, which, around 1450, initiated the wholesale replacement of an error-prone oral tradition with a written tradition and promoted widespread literacy. But printing was one of many successful communication technologies. Around 1840 Samuel Morse's telegraph used the clicking of electromagnetic relays activated by distant keys to transmit data in Morse Code. The encoding method -- dots and dashes for each letter of the alphabet -- was ideally suited for a human listening to a clicking relay. Around 1890 Alexander Bell's telephone encoded and transmitted voice signals over long-distance wires. Around 1900, Guglielmo Marconi's radio transmitted pulses (such as Morse codes) over long distance without wires; within a decade radio technologies encoded full audio signals into radio frequency (RF) signals. These technologies were extended to FM radio in the 1930s and television in the 1940s. Around 1940, the British secretly used the first electronic computers to decode German communications protected by the Enigma cipher. The first computer network of all-digital data transmission was the ARPANET in 1970; it evolved into an all-encompassing worldwide network with a billion users by 2007. The

communication and networking functions of computers are now widely accepted as more fundamental than the data processing function.

Those who say that digital computation is as revolutionary as the printing press are singling out, perhaps unfairly, one of many communication technologies. While there is no doubt that the cumulative effect of the long line of communications technologies indeed been revolutionary, not everyone is prepared to cite the computer as an agent of social revolution. James Burke, author of *The Day The Universe Changed*, says he is not ready to do this until computers can carry on natural conversations with humans.

In the early 1940s, Claude Shannon of Bell Labs developed a *Mathematical Theory of Communication* (the title of a 1948 paper and later a 1963 book with Warren Weaver) in order to lay a foundation for coding systems that could reliably overcome noisy channels. Shannon defined "information" as the number of bits required for a receiver to uniquely select the actual message sent, out of the ensemble of possible messages that could be sent. Information was thus a way of resolving uncertainty among choices. He used the thermodynamic formula for entropy to calculate the information in a message source where the probability of each possible message is known. This work, often called *information theory*, contributed greatly to the development of reliable communication technologies and established a "bits" as a measure of "information content" of a message source.

Model of Communication Systems

Shannon and Weaver offered a general model of a communication system. Every communication system contains these essential elements:

- The medium that propagates signals,
- An encoder that converts code words into signals in the medium,
- A corresponding decoder, and
- A codebook that converts received code words to messages.

They assumed that code words were represented as strings of bits. There is no loss of generality in this assumption because any symbols can be encoded into bits. The entire contents of the codebook, mapping from messages to code words, is called the "code".

This model covers all the examples cited earlier. For ship-to-ship semaphores, the channel is visual line of sight and the encoder is the sailor holding the flags in positions corresponding to letters of the

alphabet. For telegraphy, the channel is electrical copper wires and the encoder is the human operator pressing the key according to the Morse code. For telephony, the channel was originally copper wire and now includes microwave and optical fiber, and the encoder is the microphone that converts sound waves to electrical signals. For ordinary human speech, the channel is the air and the encoder is the human vocal system operated by the brain.

Noisy Channels and Error Correction

The model also recognizes noise on the channel. Noise is any influence or disturbance that can alter the signal and cause the decoder to make an error. Examples of noise: fog in ship-to-ship and smoke signal communications, accidental grounding in telegraph or telephone wires, ambient noise in human speech.

Shannon and Weaver noted that extra bits could be added to code words to compensate for noise in the channel and enable the receiver to retrieve the original code word. In 1950, Richard Hamming of Bell Labs provided a famous method to do this: a code that would correct one error per code word. Hamming's code adds 3 check bits to every 4 data bits, producing 7-bit code words, in such a way that every code word is at least three bits distant from every other code word. A single error converts the code word into a pattern one bit distant from the original code word and at least two bits distant from every other code word. The error is removed by substituting the code word one bit distant from the received pattern. His code launched a new field of error-correcting codes that complemented information theory.

Noise reduces the effective capacity (bandwidth) of a channel. To contend with (say) a 10% error rate, a communication system will have to add extra bits to the code words to compensate for bits that are lost to errors. How many extra bits is a design tradeoff:

- Add a parity bit to every 9 bits of data. A parity bit makes the total number of 1s among the 10 bits even. The receiver detects a single error when it notices that a group of 10 received bits had odd parity. On detecting an error, the receiver requests the sender to retransmit the 10 bit block. With 10% error rate, the sender will send an average of 11 blocks to ensure that 10 are accurately received. That is 110 bits total for 90 data bits, for an efficiency of 90/110 = 82%.
- Use the Hamming Code, which has efficiency 4/7 = 57%.

Minimal Codes

Shannon asked how to construct a minimal code -- the code with the fewest possible bits. This is a non-trivial question. For example, 7-bit ASCII and 8-bit ASCII encode the same text symbols, but the 7-bit code is obviously shorter. Even subtler is the possibility of a variable length code that gives the shortest code word to the most common letter -- for example, the Morse code assigns a single "dot" to the letter "e", which is the most common letter in English. In asking for a minimal code, Shannon wanted to preserve the information content of messages -- in other words, the minimum code would contain the same information as any other code, and any shorter code would lose information and not be capable of accurate reception.

Shannon's essential insight was that a code with fewer bits than the entropy could not represent all the information in the message source and thus could not guarantee that all messages could be reliably decoded. In 1952 David Huffman of MIT showed a method to construct a minimal length code whose average code word length is within 1 bit of the entropy. Huffman's code became the principle of many data compression schemes.

The Huffman code is the shortest lossless code: the original data can be fully reconstructed by reversing the code. However, if we focus on preserving only the most valuable information, instead of all the information, we can compress much further. The MP3 audio compression method, for example, deletes frequencies that are near neighbors of higher intensity frequencies -- because the ear cannot hear the fainter neighbor. MP3 can achieve compressions of 10:1 whereas Huffman coding of the same data is unlikely to do better than 2:1.

Secret Communication

It is important in many communication systems to protect against eavesdropping: a third party that monitors the channel and attempts surreptitiously to intercept and decode the signals. The science of cryptography offers principles for hiding secrets in messages. The idea is to encipher the data from the message source before it enters the sender's encoder, and decipher it when it exits the receiver's decoder. Shannon proved that a perfect cipher would result when the enciphering key is a random bit stream as long as the message: each message bit is either kept or reversed, depending on whether its corresponding key bit is 0 or 1. The input to the communication

system thus looks random and the eavesdropper can make no sense of it. The receiver can decipher by applying the same key sequence as was used to encipher. The hard part of this encryption system is getting a copy of the one-time key to both sender and receiver.

Cryptographers have devised very effective digital encryption systems that can be managed by silicon chips. The single-key systems require sender and receiver to have exchanged a (short) key in secret; they can generate enciphered sequences at rates of 100Mbs, the same as high-speed Ethernets. The two-key public systems give every receiver a public key and every sender a secret key. The two keys are related because one can decipher the other. A cryptogram enciphered under a receiver's public key can be sent by anyone but deciphered only by the receiver. A cryptogram enciphered under the sender's secret key can be deciphered by anyone (thus authenticating the sender). RSA, the most successful public-key system, derives its secret keys from the products of two large prime numbers. It relies for its security on the computational intractability of the problem of factoring a large number into two primes. A fast algorithm for factoring would become a fast algorithm for computing a secret key from a public key.

Public-key cryptosystems are much slower than single-key systems but require no advance secure key exchange. The single-key and public-key systems are often used together: the initiator of a communication generates a session key (for single-key encryption) and sends to the receiver with a public key system.

The Internet as a Communication System

In 1970, the Defense Advanced Research Projects Agency started a network of computers that exchanged digital packets rather than traditional analog telephonic signals. The network routers dynamically selected any available path to send packets to their destination, and the receiving computers reassembled the packet streams into the original messages. A system of acknowledgements and time-outs enabled the communication protocol to resend missing or corrupted packets. This ARPANET was the progenitor of the modern Internet. Its ability to dynamically reconfigure its routes made it easily extensible and its ability to resend packets made it highly error tolerant. The Internet has become the world's largest distributed communication, retrieval, and computing system. All the principles of communication cited above have contributed to its success.