

CpE/NIS 654
Speech Quality and Bit Rate

Early codecs which converted analog voice into digital bit stream for circuit switched networks (PSTN) were based on sampling the speech at 8000 times per second and each sample digitized into 8 bits, thus generating a data stream of 64Kbps. That standard for these codecs (coder/decoders) is known as G.711. In telephony, good speech quality, which obviously is based on human perception, is achieved by G.711 codecs which, as noted, generates a data stream of 64kbps. This speech quality is often referred in telephony as “Toll Quality” speech.

But with many applications requiring storing of speech in the form of data bits and transporting them over data networks as data packets, there was a great need to compress speech, yet maintain the same or near same voice quality to conserve memory, improve performance and reduce network bandwidth requirements. Thus, over the years, various algorithms have been developed and standardized which reduce the bit rates for speech to as low as 600 bps. There are systems in use today which use a maximum of 2400 bps and provide reasonably good speech quality. Speech compression has become even more important with the rapid evolution of multimedia services where speech is multiplexed as a data stream with other media and transported over data networks. VOIP service is also increasingly replacing traditional circuit switched telephony where voice is carried as data packet stream to conserve bandwidth and to take advantage of the existing data networks. Thus, a lot of work has been done on compression technologies.

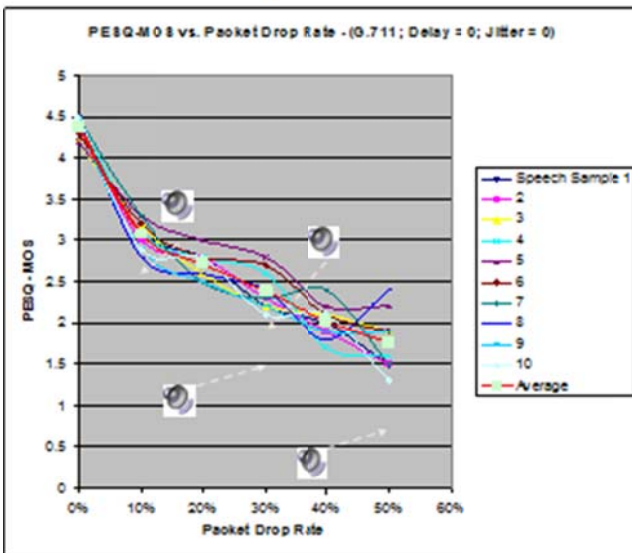
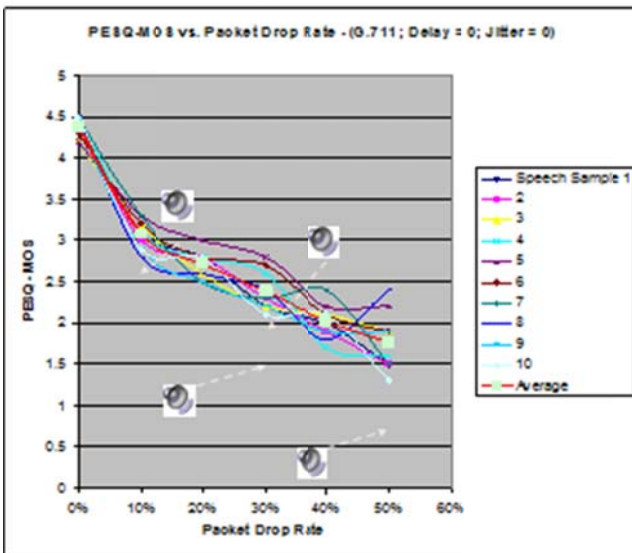
We encourage you to explore speech compression technologies and their evolution as well as how its quality is impacted when carried over the data networks. The following website illustrates the speech quality, i.e., how the speech sounds, with various levels of compression.

<http://www.speex.org/samples/>

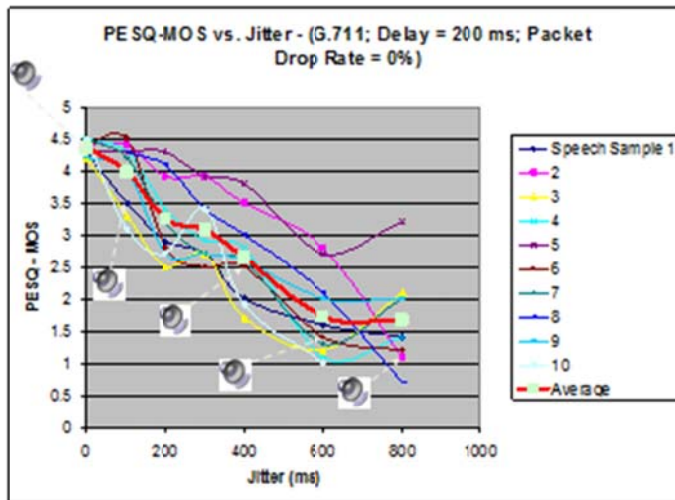
Beside the impact of compression on speech quality, it is also greatly impacted by the data network over which the packetized speech is carried. As the packets are delayed or lost, the quality of speech is adversely impacted. Attached document as well as figures below illustrates how the speech quality deteriorates as the packets are lost or arrive late. In general packets delay has less impact on speech quality than their loss or variation in delay as they arrive. Delay variations causes jitter in voice.

VOIP sample sounds.pptx

G.711 Packet Drop Rate Performance



G.711 Jitter Performance



So, how does the above discussion relates to network design? As we have discussed, key driver in network design is the performance of the applications supported by the network. This performance, in turn, is measured in terms of network capacity, delay, and RMA (Reliability, Maintainability, and Availability). Because of the special needs of speech which is very sensitive to delay and delay variation, it is important that they be understood and accounted for in the design. We encourage you to explore VOIP application, its requirements, its service architecture, and how the network design addresses this application's needs.