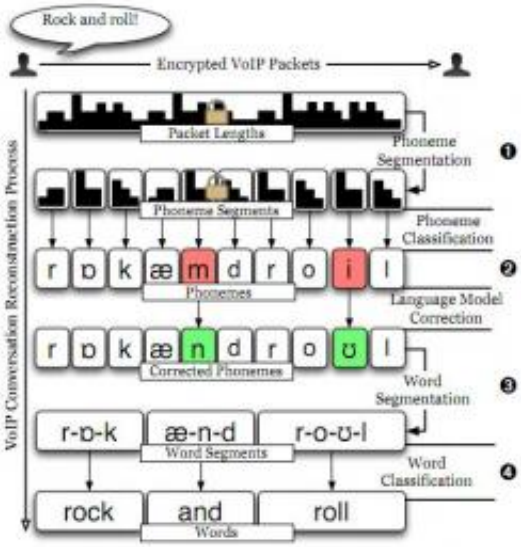
Encrypted VoIP not as secure as it sounds

May 26 2011, by Bob Yirka



Overall architecture of our approach for reconstructing transcripts of VoIP conversations from sequences of encrypted packet sizes. Image credit: Andrew M. White.

(PhysOrg.com) -- Linguistics researchers working with computer scientists at the University of North Carolina have shown that voice conversations over the Internet, even if they are encrypted, are not as secure as generally thought. Presenting their findings at the IEEE Symposium on Security and Privacy in Oakland California this past week, the team showed that by breaking up voice messages broadcast over the Internet, and then parsing the bits into phonemes (human speech components) they could, using linguistic rules, essentially recreate conversations; at least to some degree. The results varied, but were in

general good enough to gain the essence of what was being said.

The results of the team's efforts show that services such as Skype, even though they use both encoding (converting words to code or data) and encryption (transforming the encoded messages to a different form using an algorithms) techniques to prevent easy capture of voice conversations over the Internet, are vulnerable to eavesdropping by perpetuators bent on listening in on what are supposed to be private conversations.

The team was able to reconstruct conversations, not by beating the encryption scheme, but by measuring the data packet size of messages sent electronically across a network and then by applying known linguistic rules of human speech to those packets to decipher individual components of speech, which when put together, resulted in conversations that were at times, able to be understood by those listening.

In the paper that accompanied their presentation, the team describes the process as similar to that used by infants when learning to communicate. They learn by associating certain words they hear over and over with known results. When an adult speaks to them, they parse out the stuff they don't understand and instead concentrate on the words that stand out that they do know; linguists use the term "well formed" to describe terms that are understandable amongst those that are not. Infants use well formed phrases to help them deduce the meaning of other words that surround the ones they do know to try to figure out what is being said; a process the research team essentially duplicated when trying to recreate phone conversations.

Because the results varied widely, and because eavesdroppers would need a lot of time, talent and money to recreate the results the team found, current users of such services shouldn't worry that someone is listening in, but even so, now that a vulnerability has been exposed, it's

likely that Skype and other VoIP providers will take steps to eliminate the newly discovered weakness.

More information: Phonotactic Reconstruction of Encrypted VoIP Conversations: Hookt on fon-iks

Abstract

In this work, we unveil new privacy threats against Voice-over-IP (VoIP) communications. Although prior work has shown that the interaction of variable bit-rate codecs and length-preserving stream ciphers leaks information, we show that the threat is more serious than previously thought. In particular, we derive approximate transcripts of encrypted VoIP conversations by segmenting an observed packet stream into subsequences representing individual phonemes and classifying those subsequences by the phonemes they encode. Drawing on insights from the computational linguistics and speech recognition communities, we apply novel techniques for unmasking parts of the conversation. We believe our ability to do so underscores the importance of designing secure (yet efficient) ways to protect the confidentiality of VoIP conversations.