



(S) Technology that Identifies People by the Sound of Their Voices

FROM: Adolf Cusmariu
Technical Director, Content Analysis Services (S21212)
Run Date: 01/04/2006

(S//SI) "I'd know that voice anywhere..." When SIGINT transcribers work the same target set for a long time, they sometimes can identify a certain individual in recorded conversations, just by the sound of his voice and by his unique way of speaking. This process was traditionally known as "voice identification." Now, rapidly improving technology is available that can do the same job mathematically, and with it come new opportunities for the Intelligence Community. Indeed, such methods have proven surprisingly more robust and consistent than humans; refinements and improvements do continue, mainly to reduce errors and dependence on communications channels.

(S//SI) As an example of what this new "voice-matching technology" can do, state-of-the-art software and mathematical analysis verified voice intercepts from leaders of Al-Qaida. It matched speakers from various SIGINT intercepts and confirmed that certain in-theater detainees were not those suspected of terrorist activities.

(TS//SI) One specific accomplishment involved a request for an urgent voice comparison of terrorist **Abu Hakim** with an individual interrogated by the FBI in Iraq. Technical analysis showed there was no statistical reason to believe the two voices matched, a conclusion fully confirmed by a native-speaking linguist. Subsequently, the FBI arrested the individual in question based on other incriminating evidence.

(S//SI) However, transmissions by Arab broadcasters of Al-Qaida's second-in-command, **Aiman Al-Zawahiri**, showed an excellent acoustical match. Other analyses determined that the voice of **Usama Bin Laden** is unmistakable and remarkably consistent across several transmissions. During Operation IRAQI FREEDOM, it was determined that the voice claimed to be of deposed leader **Saddam Hussein** was indeed his, contrary to prevalent beliefs.

(TS//SI) In addition to supporting the Counterterrorism effort, voice-matching technologies are being applied to the emerging **Insider Threat initiative**, an attempt to catch the "spy among us." As a test, analysts mathematically compared old intercepts and audio files of NSA spy, **Ron Pelton**, who was recorded while contacting the Soviet embassy. Those comparisons showed excellent correlation. Had such technologies been available twenty years ago, early detection and apprehension could have been possible, reducing the considerable damage Pelton did to national security.

(S) Furthermore, mathematical vocal tract modeling can even identify voices largely **independently of the language spoken**. For example, a request from a customer within NSA's Counterintelligence structure posed the challenge of matching voices of a Chinese person conversing in English with two external contacts. Despite the phonetic irregularities in articulation that virtually all non-native speakers make, mathematical analysis showed nevertheless that the voices indeed match with a very high degree of confidence. Furthermore, comparison with an audio file where the person spoke Chinese with a cohort once again proved a high-confidence match.

(S) The voice-matching technology was developed by MIT/Lincoln Labs under NSA contract. It is used by S21212 in support of automated speaker identification services for a wide variety of offices within the Agency. It is also rapidly becoming the standard in the Intelligence Community.

POC: Adolf Cusmariu, S21212, Technical Director, [REDACTED]

"(U//FOUO) SIDtoday articles may not be republished or reposted outside NSANet without the consent of S0121 ([DL sid comms](#))."

DYNAMIC PAGE -- HIGHEST POSSIBLE CLASSIFICATION IS
TOP SECRET // SI / TK // REL TO USA AUS CAN GBR NZL
DERIVED FROM: NSA/CSSM 1-52, DATED 08 JAN 2007 DECLASSIFY ON: 20320108