

You are invited to an IEEE Meeting on  
Thursday, Apr 26, 2012



---

Title: **Audio Segmentation for Meetings Speech Processing**

---

Speaker: **Kofi Boakye**, Lawrence Livermore National Laboratory

Date: Thursday, April 26, 2012

Time: Presentation from 12:30 PM – 1:30 PM

Cost: No charge

Place: Livermore Valley Open Campus (LVOC) – Bldg. 6475  
Lawrence Livermore National Laboratory  
Greenville Road, Livermore, CA

Building 6475 is located south of the Eastgate Avenue entrance off of Greenville Road. Follow signs to the HPC Innovation Center. The building is located behind the UNCLE Credit Union.

RSVP: **Please make a reservation** by e-mailing Brock Beauchamp (brockb@ieee.org)

---

Meeting Description:

Perhaps more than any other domain, multiparty meetings represent a rich source of content for spoken language research and technology. Two common (and complementary) forms of meeting speech processing are automatic speech recognition (ASR)—which seeks to determine what was said—and speaker diarization—which seeks to determine who spoke when. Because of the complexity of meetings, however, such forms of processing present a number of challenges. In the case of speech recognition, crosstalk speech is often the primary source of errors for audio from the personal microphones worn by participants in various meetings. With speaker diarization, overlapped speech generates a significant number of errors for most state-of-the-art systems, which are generally unequipped to deal with this phenomenon.

In this talk I present efforts toward addressing these two issues. In both cases audio segmentation is employed to identify regions of interest—local speech for ASR and overlapped speech for speaker diarization—with a view to improving performance of the respective systems. A particular focus is the selection of features which work well for these segmentation tasks. In addition, in the case of overlapped speech I examine how two processing techniques—overlap segment labeling and exclusion—may utilize overlap information to improve speaker diarization performance.

This abstract was prepared under the auspices of the Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 (LLNL-ABS-546131)

About the Speaker:

Kofi Boakye received a B.S.E. in Electrical Engineering from Princeton University in 2002, and a Masters and Ph.D. in Electrical Engineering and Computer Sciences from the University of California, Berkeley, in 2005 and 2008, respectively. He is currently a Research Engineer in the National Security Engineering Division of Lawrence Livermore National Laboratory. His research interests include signal processing, machine learning, and computer vision.