

VoxNet - an end-to-end system to support on-line, real-time bioacoustics research

Michael P. Allen*, Lewis Girod‡, Ryan Newton‡, Travis C. Collier‡, Deborah Estrin‡

*Cogent ARC, Coventry Univ.

‡CSAIL, MIT

‡CENS, UCLA

Introduction: Processing challenges in bioacoustics research

Studying acoustic communication

- **Dynamics of acoustic communication**
Bioacoustics research field has interest in understanding behavior of animals/birds through vocalizations; e.g. *marmots*, *antbirds* and *wolves*
- **Three important challenges**
Census, *classification* and *localization* are three important aspects of studying animal or bird behavior from vocalizations
- **Data processing**
Typically these observations are made manually or using suitably configured *automated event detectors* which pick out events of interest for further processing

Data gathering/processing approaches

- **Data gathering**
 - Scientists would traditionally use *manual observation* or an *array of wired microphones* to record acoustic data over area of interest
 - Also, Acoustic ENSBox network (IPSN 2007) allows wireless, time synced data gathering, OR *on-line automated event detection*
- **Data processing**
 - Wide variety of standalone/integrated tools – no one specific approach
 - Matlab, Labview, Audacity, Sox, Baudline, Raven, Ishmael, RTS/SIGNAL

Problem: Using a purely off-line or on-line approach can be problematic

Off-line processing

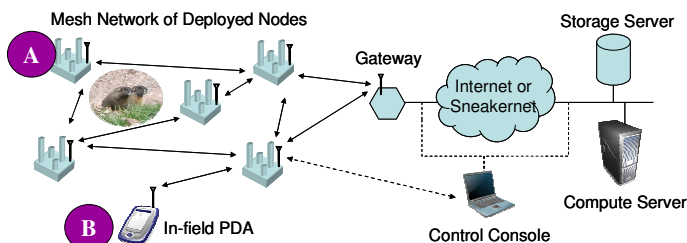
- Scientist cannot interact with data processing/gathering in the field
- Potentially can miss useful interactions based on feedback of system (i.e. take a photo based on localization results)
- Current off-line processing tools vary in generality and complexity
- Using more powerful recording tools creates very large datasets – for instance, 8 hours continuous recording @ 48KHz = 10GB/Node (typically 8 Acoustic ENSBoxes/deployment)

On-line processing

- Potentially huge data sets can be reduced as data is recorded to make processing more manageable, using on-line event detectors
- However, reducing data set may reduce its usability – badly configured detectors may yield bad results (missed detections for example, which is problematic for census)
- The data lost may have been useful for another purpose, for example to identify other animals/birds present in recordings

VoxNet: A system to support bioacoustics research both on-line and off-line

Overview of proposed architecture



Hardware for distributed acoustic sensing

- **Existing Acoustic ENSBox**
 - 2x400 MHz PXA processors, 64 MB RAM, 16GB CF
 - 802.11 wireless, 4 x 48KHz audio chans
 - 20.8 x 14.4 x 9.2 cm, 2.3 kg, 7.5W consumption
 - Supports time sync, self-localization
- **Proposal for new h/w platform (VoxBox)**
 - 1x600 MHz PXA processor, 128 MB RAM, 16GB CF
 - 10x increase in energy efficiency, 20x comms. Range
 - 19 x 14.5 x 5.4 cm, 0.75 kg, 0.75 W consumption



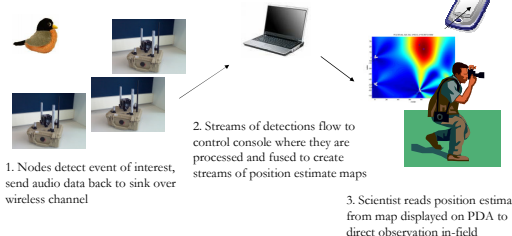
Queries as Applications

- Nodes have a set of 'core' functionality – self-localization, time sync, routing, archiving
- Actual 'applications' are queries, which are programmed using Wavescript (MIT)
- Wavescript takes macro-programming approach, abstracting out networking issues for the programmer
- Queries can be run over offline data also

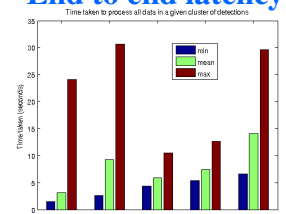
Dynamic reconfigurability

- Queries can be dynamically replaced during system run-time, to allow most flexibility
- Queries are compiled at the Control console, and disseminated over reliable publish/subscribe streams to all nodes simultaneously
- Performance is comparable to manually copying binary files to each node in turn

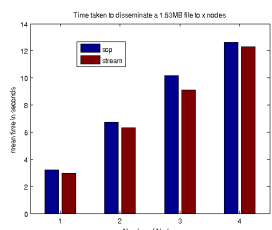
An example Query



End to end latency



The time taken to process a cluster of detections corresponding to a single vocalization. Graph was re-created off-line from data log using a different clustering algorithm than used in the field (experimentation from RMBL, CO).



Mean time taken to transfer a 1.5MB file to a certain number of nodes over one hop. A bar is the mean of 5 tries.

Visualization of data

- Using PDA in-field, scientist can 'subscribe' to streams which run in the network (node, sink, network), and visualize the data that is coming from them (see example Query above)
- Visualizers can be dynamically plugged into data streams – for example, a time domain visualizer for audio data, or a polar plot visualiser for DoAs