Using Sensor Networks to Classify Frogs Based on Their Calls

Eduardo Freire Nakamura
Research and Technological Innovation Center (FUCAPI)
Federal University of Amazonas (UFAM)

Thursday, May 24, 2012
Research Challenges

Grand aligned challenges

Ubiquitous Computing (1)
Biological Diversity and Ecosystem Functioning (2)

We need to understand the Earth’s physical systems
Climates, geology, hydrology, ...

The rainforest is a key environment
Let’s start with the Amazon forest

Key Problems

An environment of extremes
Relative humidity: 70-90%
Temperature: 64-122 °F
Huge area, limited accessibility

Additional challenges to WSNs
Ecological issues
Where should we start?
The Anura Project

ANURA: Sensor Networks for Classifying and Monitoring Frogs Based on Their Vocalizations as an Early Indicator for Ecological Stress in Rain Forests

Financial Support
LACCIR - Microsoft
PRONEX - FAPEAM/CNPq (Brazil)
Motivation

Amphibians
Amphibians are very sensitive to changes (Carey et al., 2001)

Anurans (Frogs and Toads)
Closely related to the ecosystem (Alexander & Eischeid, 2001)
Fairly easy to be monitored
Our Approach
Data Gathering
### Data Gathering

#### Adenomera Andreae

<table>
<thead>
<tr>
<th>Species</th>
<th>Individuals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hylaedactylus</td>
<td>8</td>
</tr>
<tr>
<td>Rhinella granulosa</td>
<td>3</td>
</tr>
<tr>
<td>Adenomera andreae</td>
<td>8</td>
</tr>
<tr>
<td>Ameerega trivittata</td>
<td>5</td>
</tr>
<tr>
<td>Hyla minuta</td>
<td>11</td>
</tr>
<tr>
<td>Hypsiboas cinerascens</td>
<td>2</td>
</tr>
<tr>
<td>Leptodactylus fuscus</td>
<td>4</td>
</tr>
<tr>
<td>Osteocephalus oophagus</td>
<td>4</td>
</tr>
<tr>
<td>Scinax ruber</td>
<td>4</td>
</tr>
</tbody>
</table>

**Total**: 49
Classifying the Anuran Calls

Preprocessing

Syllable segmentation

Preemphasis

Feature Extraction

Classification

kNN

SVM
Feature Extraction

Time domain

(R) Zero Crossing Rate
(S) Spectral Centroid
(B) Bandwidth

Frequency domain

MFCC
*Mel-Fourier Cepstral Coefficients*
### Best Results

<table>
<thead>
<tr>
<th>Features Used for Classification</th>
<th>5-NN</th>
<th>10-NN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\alpha = 0.4$</td>
<td>$\alpha = 0.5$</td>
</tr>
<tr>
<td>MFCC (animals method)</td>
<td>97.07%</td>
<td>97.12%</td>
</tr>
<tr>
<td>RSB (anura method)</td>
<td>87.98%</td>
<td>89.83%</td>
</tr>
<tr>
<td>R &amp; MFCC</td>
<td>97.27%</td>
<td>97.40%</td>
</tr>
<tr>
<td>S &amp; MFCC</td>
<td>98.10%</td>
<td>98.14%</td>
</tr>
<tr>
<td>B &amp; MFCC</td>
<td>97.75%</td>
<td>97.91%</td>
</tr>
<tr>
<td>RSB &amp; MFCC</td>
<td>98.41%</td>
<td>98.53%</td>
</tr>
</tbody>
</table>
Towards Efficient Data Gathering

Compressive Sensing
Reconstruct the signal based on samples
“Generalization” of Fourier and Wavelet Transforms

- Fast Fourier Transform
- Fundamental Frequencies
- Audio Reconstruction Base
- Audio Samples
- Gradient Projection for Sparse Reconstruction
- Reconstructed Signal
Some Results

1.8 MB of sound

- Classification Rate (%)
- Packet size (KB)
- Compression parameter ($\alpha$)

94%
$\alpha = 45$
20KB

Classification Rate (%)

Compression parameter ($\alpha$)
Some Results

Graph 1: Network Energy Consumption (μAh) vs. Anuran Vocalization Range (m)
- 45 KB (α = 15)
- 1.8 MB (α = 1)

Graph 2: Average Delay (s) vs. Anuran Vocalization Range (m)
- 45 KB (α = 15)
- 1.8 MB (α = 1)
Data Clustering and Node Clustering

Improve Classification (data clustering)
Reduce energy consumption (node clustering)

First Approach
K-Means for data clustering
Our Approach
Our Approach
Our Approach

Value Fusion
Data Averaging

Local Decision

Value Fusion
Data Averaging

Local Decision

Local Decision
Our Approach

Value Fusion
Data Averaging

Local Decision

Decision Fusion
Voting

Value Fusion
Data Averaging

Local Decision

Local Decision
What we discovered

The best value for k
Positive correlation with the number of species

Finding the best k
Tell us how many species are calling
Conclusions

Frog calls
Can be used for classifying species
Still cannot classify individuals

Compressive Sensing
Efficient Data Gathering
Saving 88% of energy consumption (90% of data compression)
Classification success is close to 95%

Data clustering
Give us a hint on the number of species
Even when we do not know the species
Desired Deployment Site
Acknowledgement

Our Research Network

Prof. Eduardo Nakamura (FUCAPI/UFAM, Brazil)
Prof. Edgar Vallejo (ITESM, Mexico)
Prof. Antonio Loureiro (UFMG, Brazil)
Prof. Alejandro Frery (UFAL, Brazil)

Our Local Team

Prof. Eulanda dos Santos (UFAM)
Prof. Fabiola Nakamura (UFAM)
Prof. Horácio Oliveira (UFAM)
Prof. Marcelo Gordo (UFAM)
Prof. Maurício Figueiredo (FUCAPI)
Ph.D. Student Efren Souza (UFAM)
M.Sc. Student Juan Colonna (UFAM)
M.Sc. Student Andre Campos (UFAM)
M.Sc. Student Antonio Ramos (FUCAPI)
M.Sc. Student Afonso Ribas (FUCAPI)