Poster : *Resource-Efficient Detection of Elephant Rumbles*

Namal Jayasuriya, Tharindu Ranathunga, Kasun Gunawardana, Chamath Silva
Prabash Kumarasinghe, Asanka Sayakkara, Chamath Keppitiyagama, Kasun De Zoysa

University of Colombo, School of Computing, Sri Lanka
nmj,kgg,jpk,asa,chamath,kasun[@ucsc.cmb.ac.lk],tharindu@scorelab.org, mbckchamathsilva@gmail.com

Kasun Hewage, Thiemo Voigt
Uppsala University, Sweden
kasun.hewage@it.uu.se, thiemo@sics.se

ABSTRACT

The human-elephant conflict is causing significant damages to the life of human and elephants in Sri Lanka. Minimizing encounters between humans and elephants is hence crucial in alleviating the human-elephant conflict. We have earlier introduced Eloc, a cost-effective localization system that is based on infrasonic emissions from wild elephants. One of the remaining challenges is a resource-efficient method to detect the elephants' infrasonic emissions. We present our efforts on devising a support vector machine (SVM) that is able to detect elephant rumbles on Eloc nodes.

1. INTRODUCTION

The human elephant conflict has been a profound problem in Sri Lanka for decades that has caused the loss of a significant number of both human and elephant lives. Increasing human settlements and cultivations close to forest areas has resulted in shrinking elephant habitats and blocking their migration paths. In 2010, 81 people were killed by elephant attacks [3] while every year about 200 elephants die as a result of collisions [3]. At present, the primary solution is the use of electric fences around elephant habitats to prevent elephants venturing beyond their habitats. This is an expensive and potentially life threatening solution.

As an alternative solution we have proposed Eloc [5], a cost-effective system to localize wild elephants using their rumbles. An elephant rumble is an infrasonic sound emitted for the purpose of communication [4]. Eloc uses the rumbles to determine the location of the sound source and hence localize elephants. Since infrasonic sound travels long distances without significant attenuation, Eloc is able to detect elephants from large distances. Eloc consists of a number of acoustic sensor nodes deployed across the outskirts of human settlements which border elephant habitats (see Figure 1).

One missing piece in our current solution is a way to detect elephant rumbles in the sound recordings. Due to the noise in the sound recordings stemming from wind, vehicles and other sources, detecting elephant rumbles is challenging. Furthermore, we would like to perform the detection on the Eloc nodes themselves rather than at a powerful server in order to save valuable bandwidth required for transferring the recordings. This is of particular importance since our system requires rumbles on recordings from at least two sensors which increases the bandwidth requirements. Therefore, it is important to detect the elephant rumbles at the Eloc nodes themselves and transmit data only when a rumble is detected. Eloc nodes are, however, resource-constrained comprising a 1GHz single core microprocessor and 1GB of RAM.

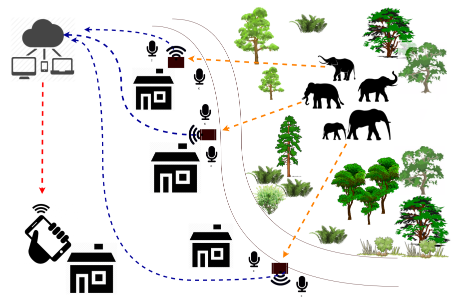

Figure 1: Application Scenario

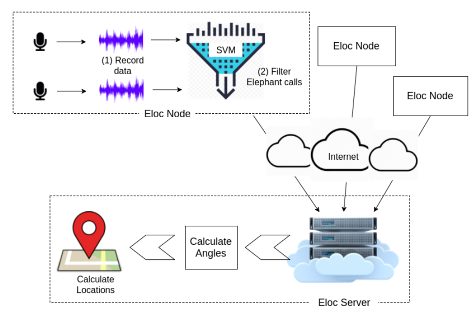

Figure 2: Design Overview

2. DETECTING RUMBLES ON ELOC NODES

Figure 2 presents the high-level design of the elephant rumble detection in the context of the Eloc system. After the recording of acoustic data, the eloc nodes separate unwanted noise from the part of the recordings that contain

SenSys '17, November 6–8, 2017, Delft, Netherlands

DOI: https://doi.org/10.1145/3131672.3136982

elephant rumbles. After the filtering, Eloc nodes transmit those recordings to the Eloc server, where the angles and hence the elephant locations are computed.

For detecting recordings with rumbles, we take a machine learning approach since our problem is a typical binary classification problem: either there are rumbles in a recording or not. Furthermore, we have a set of annotated recordings which allows us to apply supervised learning techniques. We opt for a support vector machine (SVM) since it is a well-known supervised learning techniques commonly used in binary classification problems and provides robustness in class border demarcation. Eloc nodes feature a mini processing unit which comprises 1GHz single core microprocessor and 1GB of RAM. In order to cope with these restrictions we perform the learning offline and run the trained SVM classifier on Eloc nodes.

In order to train the classifier, we train the SVM with samples from both classes, with and without elephant rumbles. Since inputs to a SVM are drawn from a multidimensional feature space, it is necessary to transform sampled acoustic data into a relevant feature space. Therefore, we adopt the Mel Frequency Cepstral Coefficient (MFCC) for feature extraction. MFCC is a widely used audio feature extraction method from the human speech processing domain [1] to transform these raw signals into a corresponding feature space. However, to fit with the features of infrasonic fundamentals and their lower harmonics of elephant vocals, we shift the frequency range of the MFCC to 14Hz to 100 Hz. We set the sampling window size to 0.25 seconds to make sure that several wavelengths of the maximum possible wavelength in an elephant rumble are included in a single sample window. Further, we shift the sampling window by 0.1 seconds which is equal to the time that is taken by the lowest frequency of an elephant rumble (10Hz). Hence, extracted MFCC feature vector for each window represents an input sample to the SVM.

3. EVALUATION

In our evaluation we have used 149 audio recordings of Elephant rumbles that have been classified by domain experts [2]. The data set has been recorded at a sampling rate of 44100 samples per second. However, since the ADC in the pcDuino used for the Eloc node cannot take samples at that rate, we down-sampled the original data to 8000 Hz to match capabilities of the single board computer. Then these data sets were transformed into the corresponding feature space using our modified MFCC feature extraction.

We use 70% of data for traing, 30% of data for testing out of 10426 positve windows and 10736 negative windows. The classification accuracy was reported as 79%. It Typically, a rumble spans longer than a half a second. Therefore, to further increase the accuracy we decided to consider consecutive five windows of a recorded sound clip as the elephant detection criteria. This modified mechanism resulted the accuracy of 93%.

Further we investigated the detection time of a trained SVM model for different audio lengths and it revealed that the detection time linearly increases with the audio length (Figure 3).

The initial objective of this research was to reduce the data to be transmitted over the network by introducing a node level filtering mechanism to the Eloc node. The preliminary results in the study shows the applicability of SVM technique to achieve this objective on low cost hardware infrastructure. The Eloc node can only store a sample of about 2 seconds duration due to the limited memory available on the pcDuino. As shown in Figure 3 it takes about half a second to classify 2 seconds of data and we transmit the data over the network only if an elephant rumble is detected. The significant detection accuracy we obtained clearly indicates this approach will help to decrease the data to be transferred by a single Eloc node which ultimately reduces the network traffic.

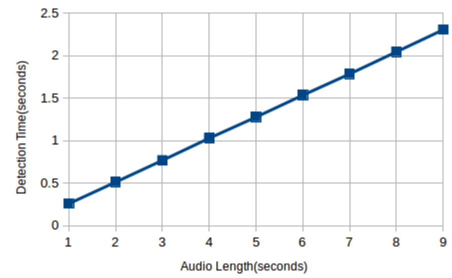

Figure 3: Classification time for different lengths of sound clips on the Eloc node (using 8000Hz samples).

4. FURTHER DEVELOPMENTS

There is still room to improve the reliability and accuracy of the system. Note that when the Eloc node is running the classifier it does not collect data samples. Therefore, it is important to further reduce the classification time without compromising the accuracy. We also want to bring down the sampling rate to reduce the network traffic and the amount of data collected on the Eloc node. However, we noted that the classification accuracy also decreases with the reduction in the sampling rate which makes this a challenging task.

5. REFERENCES

[1] S. Davis and P. Mermelstein. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE transactions on acoustics, speech, and signal processing*, 28(4):357–366, 1980.

[2] S. de Silva. Acoustic communication in the asian elephant, elephas maximus maximus. *Behaviour*, 147(7):825–852, 2010.

[3] P. Fernando, J. Jayewardene, T. Prasad, W. Hendavitharana, and J. Pastorini. Current status of asian elephants in sri lanka. *Gajah*, 35:93–103, 2011.

[4] K. B. Payne, W. R. Langbauer, and E. M. Thomas. Infrasonic calls of the asian elephant (elephas maximus). *Behavioral Ecology and Sociobiology*, 18(4):297–301, 1986.

[5] A. Sayakkara, K. Hewage, T. Voigt, N. Jayasuriya, T. Ranathunga, N. Vithanage, C. Keppitiyagama, and K. De Zoysa. Eloc: Locating wild elephants using low-cost infrasonic detectors. In *Distributed Computing in Sensor Systems (DCOSS), 2017 International Conference on*. IEEE, 2017.