

Bat Banter: Using Unsupervised Graph-Based Clustering to Discover “Phrases” in Bat Communication



Melinda Repko (nps.gov)

LINK TO COLAB CODE:

<https://colab.research.google.com/drive/1d0bTmW6VmKTNDHwQ2Kapj4aFPXuaZ5ul?usp=sharing>

[1] Introduction

Bat experts (or chiropterologists, if you're being fancy) have known for decades that the ultrasound “chirps” that bats emit serve purposes far beyond just echolocation or honing in on mosquitos, moths, and other tasty morsels. Bats also use their chirps to communicate with other bats, and biologists have found that they talk about subjects as wide ranging as roosting disputes (think, “Hey, I was sleeping here!”) to mating chirps. These analysis have been done by training clustering algorithms on individual chirps which have been labeled using videos of bat interactions and labeling particular chirps with contexts like food disputes, roosting space disputes, and mating (Prat et al. 2016).

There has also been analysis done of “phrases” within bat communication, such as a 2023 study from the same lab which used annotated sequences of chirps to train a classifier which labeled sequences as belonging to those same contexts of food, space, and mating (Amir et al. 2023). This study presented a model which classified sequences into the correct contexts with around 65% accuracy, and that permuting the order of chirps within a phrase did not affect the classifier’s accuracy, with the implication that the order of chirps does not contribute to contextual information. However, they did not test whether reordered phrases actually occur in bat communications (e.g., a classifier could probably guess that “I made chicken for dinner” and “Dinner for chicken made I” are both about food, but only one appears in regular English communication). Determining which phrases occur frequently in bat communication is important to biologists who are collecting terabytes worth of bat chirp recordings, but don’t know much about what exactly they should be looking for.

The goal of my project is to develop a graph-based, unsupervised clustering approach to identify repeated structures in bat communication. This type of work would be of most value to biologists, who often don’t have the luxury of being able to associate the bat recordings they collect with videos, as videos are more expensive and less practical to collect. Unsupervised analysis is also the only feasible method for analyzing night behavior, as bats are feeding and moving around large outdoor areas, and are therefore nearly impossible to video. Graph-based analysis will allow me to emphasize the temporal connection between chirps, in order to identify phrases in orderings which occur in the wild, and perhaps even identify groups of phrases which often occur in response to one another. Once biologists have a set of common phrases, they could identify those phrases across different locations, times, and perhaps even species, allowing them to identify more complex patterns in large-scale bat behavior.

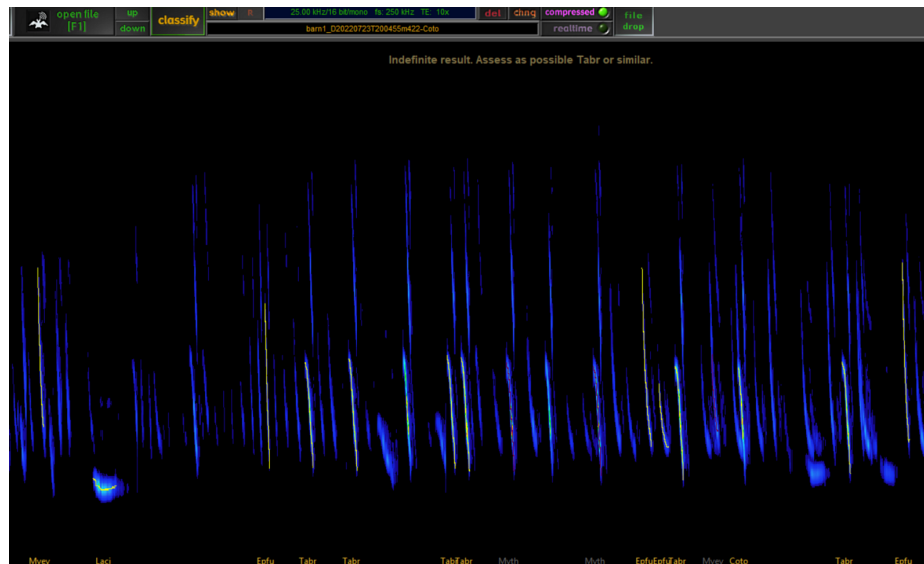
. . .

[2] Data Processing

[2.1] Preliminary Data Analysis

The raw data came in the form of a series of .wav files of bat chirps from the Jasper Ridge Biological Preserve, recorded by a microphone stationed by a small lake on the preserve. [more info about data]

To convert the raw data into usable features, I used a software called SonoBat 30 which is made for bat call analysis. SonoBat performs species classification and also can export more than 100 acoustic features for each chirp.



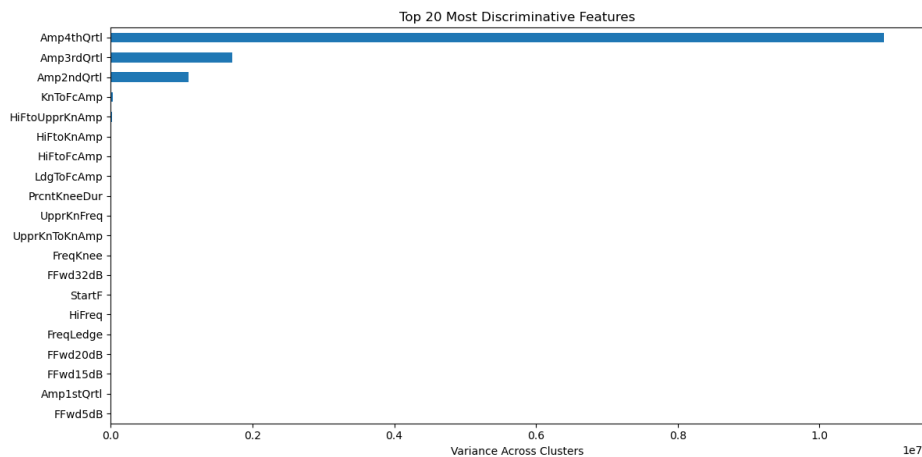
A SonoBat visualization of the lake bat chirps

I filtered my dataset to only include the species commonly known as Townsend's Long-Eared Bat as a simplifying factor so that the model would not be trying to predict across species. This left me with around 116,000 chirps, each with 120 features. I performed PCA on the features and reduced them to 30 components, explaining 90.22% of the data.

I started with individual chirp classification, because the Amir et. al study found that when classified individually, 80% of chirps fell into the same context classification as the phrase they were in, implying that chirps themselves communicate contextual information. I was interested in whether clusters of individual chirps would mirror the similar structures to clusters of phrases.

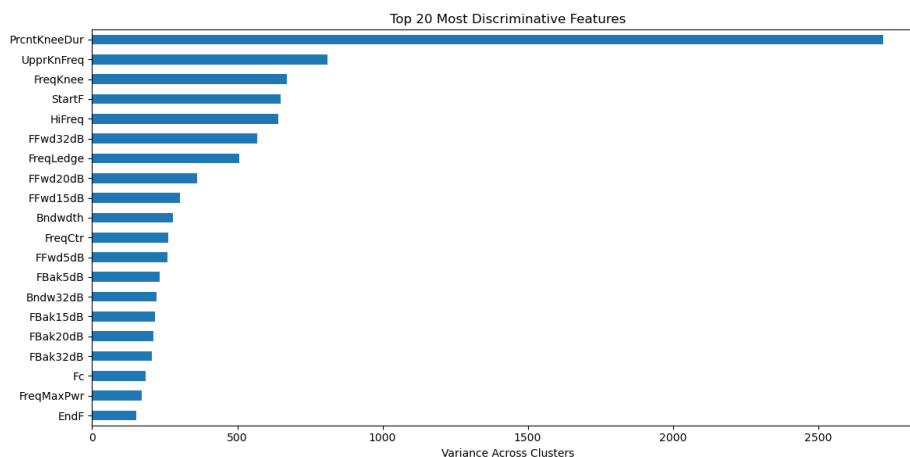
To gain an initial understanding of my data, I performed individual chirp clustering using HBDSCAN. This yielded a total of five clusters,

but I noticed something concerning about the features which most explained the differences between the clusters. The variance was dominated by three features: “Amp4thQrtl,” “Amp3rdQrtl,” and “Amp2ndQrtl,” representing the amplitudes across different sections of the chirp.



Initial discriminative features for individual chirp clusters, before removal of amplitude features.

This analysis helped me realize that in my scenario, amplitude was not an indicator of the underlying acoustic properties of the chirp as much as it was an indicator of how close a bat was to the microphone when it emitted the chirp. To solve this issue, I removed all features related to amplitude except those which captured relative amplitude information within a chirp, like “AmpVariance.” After running clustering again without the amplitude features, the most discriminative features looked much more promising. [More info about final clusters here].



Most discriminative features for individual chirp clustering after removing amplitude features

After removing features related to amplitude and metadata which wasn't relevant to the acoustic properties of the chirps, I was left with 75 features for each of the 116,000 chirps. I re-ran PCA, yielding 30 components which now explained 94.67% of the variance.

[2.2] Data Processing for GNNs

I first define each chirp as a node with 30 features, represented by the 30 PCA components. There are a number of individual recordings which make up the entire set of 116,000 chirps. These recordings are represented by the feature "file_id", which I removed for PCA but used to split the chirps into sub-graphs, one for each recording. Within each subgraph, consecutive chirps are linked by an edge with a single feature representing the time interval between the chirps. The structure of each subgraph is therefore a 1-dimensional chain of chirps, similar to how a sentences would be represented if we think of each chirp as one word.

. . .

[3] Models/Methods

[3.1] GraphSage Design & HDBSCAN(Baseline)

[3.2] Graph Attention Network (GAT) & HDBSCAN

[3.2.1] GAT Model Design & Metrics

I performed initial analysis by setting the sequence length k to 4, and grouping the chirps into non-overlapping sequences. Sequences are composed of directly temporally adjacent chirps and do not contain chirps across multiple recordings.

I then create contextualized 32-dim embeddings for each chirp using a 2-layer GAT, then do average pooling to create sequence-level embeddings. Finally, I use HDBSCAN to cluster based on the sequence-level embeddings.

[3.2.2] GAT Evaluation

(not yet implemented in code)

To evaluate the clusters, I will calculate the Silhouette score, which measures the separation of clusters vs. their tightness, and will also calculate the Calinski–Harabasz Index, which measures the variance ratio across clusters. To evaluate cluster stability, I will perform clustering on multiple random sub-samples of 50,000 chirps and compare the resulting clusters. The variation of information between the clusters found in each subsample should be low, below some metric which I have not yet determined

[3.3] Graph Attention Network (GAT) & Deep Modularity Networks (DMoN)

As an alternate method to learning embeddings with GAT and then running a separate clustering algorithm, I will introduce clustering directly into the GAT architecture using a Deep Modularity Network architecture, an unsupervised graph pooling method published in 2023 (Tsitsulin et. al, 2023).

. . .

[4] Results

. . .

[5] Conclusions

. . .

Bibliography

Amir et al (2023):

<https://www.sciencedirect.com/science/article/pii/S2589004223005436?>

Tsitsulin et al. (2023): <https://jmlr.org/papers/volume24/20-998/20-998.pdf?>

Prat et al (2016): <https://www.nature.com/articles/srep39419?>

