

Aircraft Environmental Noise Contamination Detector Proposal

Rahul Birmiwal
<e-mail>

Sean Miller
millsea0@uw.edu

Todd Schultz
tschultz@uw.edu

Executive Summary

A proposal is presented to assess the feasibility of automated machine learning-based algorithms to replace human listeners for environmental noise contamination detection for noise certification fly-over testing. The potential business impact could result in large cost saving with the reduction of testing personal, equipment, and improved testing efficacy. This proposal leverages decades of research work using acoustic signal features and classification methods from other industries to compose a list of features and classifications to propose here. The proposed plan is to explore a wide range of feature selection options such as octave spectrum, wavelets and with a wide range of classifiers such as logistic regression, neural networks, and long short time memory networks to understand the performance of each and to recommend the best performers to move forward into additional studies.

Project Background

The impact of aircraft noise on a “community” is well-regulated by the Federal Aviation Administration [1] (FAA). Aircraft manufacturers such as Airbus and Boeing must adhere to these standards for all new or derivative aircraft models which includes flight testing of prototype aircraft. Community noise flight tests consist of instrumenting one end of a runway of a remote airport with acoustic recording devices and recording the acoustic signatures as the test plane is flown over the instrumentation for the many conditions as required by the Federal regulations. To avoid non-compliance or costly and time-consuming repetitions of the tests, the surrounding acoustic environment should cause no noise contamination on the recordings that could invalidate each test. Sources of contamination can include but are not limited to: bird chirps, wildlife/livestock vocalizations, insect noises, traffic noises, and aircraft noises borne from aircraft besides the target test flight. Current testing procedures use extensive equipment and human labor to detect, assess, and remedy any environmental noise contamination to ensure the recorded acoustic signatures are solely from test aircraft.

This project is a feasibility study into using automated procedures for detecting and assessing the presence of environmental noise contamination in an audio recording. The goal would be to extend this research into a real-time system to monitor, detect, and classify the presence of environmental noise. The scope of the classification should be sufficient to guide test engineers on the necessary actions to take to remediate the contaminated data. For example, whether the flight condition need be redone or what type(s) of contaminant noise sources need to be removed from the testing site. Human labor designated for these tasks, while reasonably effective, is taxing and not cost-efficient. Thus, a proposed signal processing and machine learning-based system for this problem.

Target Audience

The target audience of the noise contamination detector is an aircraft manufacturer, the sponsor Boeing Test & Evaluation (BT&E). This project could remove the need for multiple on-site work stations and operators and could lead to significant cost reductions for BT&E. The performance of the system could increase the accuracy and repeatability of detection, further streamlining regulatory noise testing by BT&E.

Constraints and Requirements

The project, its models, its software code, and data must not limit the commercial use by The Boeing Company. Furthermore, all software code must be executable in MATLAB, the preferred computational platform for the BT&E noise testing division.

Information Required from the Sponsor

Acoustic signal training data that is labeled with the presence of contamination and what the contaminating noise is if it exists.

Project Proposal

This project seeks to research, develop, and implement a signal processing into machine-learning pipeline for automating the detection of environmental noise contamination contained in acoustic measurements. A proposed methodology for achieving such results could be:

Feature Selection

This portion of the project will require research and implementation of novel audio signal processing techniques to create and select signal features relevant for the classification task. Signal filtering, Fourier analysis, signal-to-noise ratio/signal energy spectrum methods, wavelet transforms, and other time-series transient signal classification methods are potential methods for this purpose. Further exploration in which techniques are likely to be prioritized in the section for Current Research and Work in this field.

Model Training

Given a set of covariates from feature selection the project will train classification algorithms using the training data provided by BT&E. These machine-learning (ML) algorithms will range from basic to complex and are likely to encompass everything from softmax multi-classification and multi-layer perceptrons, to recurrent neural networks. The algorithms will be considered for their pre-trained execution time as the intended goal is implementation into a real-time monitoring system. Choices for models that we will explore are also enumerated in the Current Research and Work section.

Model Evaluation

Feature selection and model training will be continuously re-evaluated according to model performance. It is plausible that the features engineered will not be sufficient to result in accurate classification and different signal processing techniques must be researched and deployed.

Data Augmentation

To assist in the prior steps, the data set from the sponsor may be augmented with other examples of environmental noise that allow for unrestricted commercial use, such as data from the United States National Park Service [2].

Once satisfied with the performance of the pipeline, robust documentation to retrain the algorithm(s) and reproduce and/or adapt this project for BT&E will be generated.

Project Deliverables

Trained Model

The primary deliverable of this project is an algorithm trained on the provided dataset. Along with this model this project will providing documentation explaining its selection and theoretical background for detecting audio signals contaminated by environmental noise sources. Measures of the algorithm's performance for a wide spectrum of noisy audio signals and contamination sources will also be provided.

Given the constraints of this project, the algorithm will be written and executable in MATLAB. This information will give BT&E the necessary background from our research to understand why the model was chosen over other considered models/feature sets and set expectations around performance of the algorithm once implemented within their own systems.

Documentation

In addition to the model, this project will provide documentation covering the process for algorithm retraining and overall project reproducibility. This information is necessary for the sponsor to understand how to update the model with new datasets.

Other Explored Models/Feature Sets

Our final deliverable will be a document enumerating the algorithms investigated during the course of this project. Included in this will be an explanation of each one's limitations and advantages. This document will provide the information to reinforce our suggested implementation and give BT&E the context for our decision.

Project Software

- MATLAB
 - o Audio System Toolbox
 - o Deep Learning Toolbox
- Trello (Project Management)
- Git/GitHub

Previous Work

A variety of features and techniques are used in modern research to identify acoustic signals. Much of the early work started with spectrum-based features from either constant percentage bandwidth or octave spectrum or discrete Fourier transform spectrum. Later work has been expanding the list of features and algorithms to more advanced and exotic methods. Recently, Wang et al (2014) proposed an improved time encoded signal processing algorithm coupled with a machine learning classifier to identify different vehicle types from acoustic or seismic signals. Earlier, Wu et al (1998) applied principal component analysis to discrete Fourier Transform spectrums to generate features for acoustic vehicle recognition. Andén and Mallat (2014) investigated and compared many options for feature generation and classification such as spectrograms (short time Fourier transforms), Mel-frequency cepstral coefficients, and wavelet scattering transforms for musical genre classification. Wavelet package have also been investigated for acoustic bearing fault detection in Hemmati et al (2016).

In Adavanne, Parascandolo, Petila, Heittola, Virtanen, 2016, LSTM units (Long short-term memory) within a Recurrent Neural Network (RNN) are leveraged to identify sound events in polyphonic audio samples [6]. Three different feature sets were studied in their evaluation, log mel-band energy, harmonic features (such as pitch), and time difference of arrival (TDOA) and found that a combination of mel-band and TDOA features performed the best. A preliminary survey of the data for this project has already seen some success in leveraging harmonic features by visualizing the octave spectrum from contaminated and non-contaminated audio.

One technique that considered but decided against, due to the small size of the dataset was metric or similarity learning. Royo-Letelier, Hennequin, Tran, Moussallam, (2018) explored the use of metric learning to disambiguate artists within a music catalog. They found that at smaller dataset sizes (300 or fewer examples), that a traditional 1D-CNN (One Dimensional Convolutional Neural Network) outperformed the metric learning model. The model they compared their metric learning model to is described in Park, Lee, Park, Ha, Nam (2017). An idea that the project may explore given enough time is leveraging triplet loss as they do in this paper to separate contaminated and non-contaminated audio. The goal of triplet loss is that given a tuple of (x_a, x_+, x_-) to learn a function F where the similarity of x_a to x_+ is greater than that of x_a to x_- .

Appendix A - Data Pipeline

The data used for this project originates from files. Typically, wave binary files, but the input data sources are not restricted to only this filetype. Since this is a research and development project, all data is expected to be in file format instead of streams or API calls. Additional sample data may be required to expand the examples of wildlife noises.

The data set consists of 67 professionally recorded and labeled audio files near a major international airport. Figure 1 shows the installation of the microphone in the field with a hemispherical wind screen to mitigate wind noise. The data were recorded on three separate days in the months of November 2017, December 2017, and January 2018. The data consist of 49 aircraft recordings, 7 wildlife noise recordings, and 10 ambient recordings and vary in length from 10 seconds to several minutes. Each recording was captured with scientific instrumentation grade equipment at a sampling frequency of 51.2 kHz and with a usable bandwidth from 4.6 Hz to 20 kHz. No missing samples or other anomalies are present in the data; however, the recorded levels have been normalized to allow for public release. This should not impact the project as the levels are to be normalized to remove the effects of differences in propagation distances from sound sources to the microphone sensors. Examples for the time series recordings for a bird chirp and a jet aircraft are shown in **Figure 2** and **Error! Reference source not found.** respectively.



Figure 1: Microphone installation

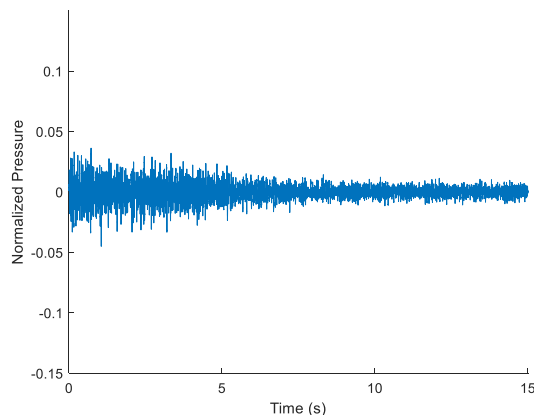


Figure 2: Example of the time series data for a bird chirp. (From file bird01.wav)

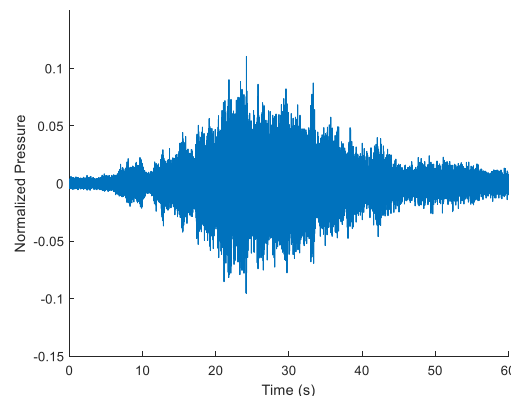


Figure 3: Example of the time series data for a jet aircraft. (From file plane08.wav.)

Data Issues

Despite the data being of high quality, there are concerns the project should be mindful of when performing this work. The recorded data is heavily skewed towards aircraft signatures thus creating an imbalance in the number of samples for each class. Additionally, each signal can produce 50-100 blocks or samples of features for classification and each block may contain 10's to 1,000's of features depending on the how features are generated. This will present a data management concern as all the data, features, and classification labels must be accurately tracked through the data processing. The class imbalance issue can be addressed by creating additional data by combining a clean aircraft signature with a contamination signal. This has the added benefit of being able to control and study the impact that the signal-to-noise ratio has on the system performance. If these combinations aren't sufficient, public

domain recordings are available for use to extend the set of wildlife/livestock vocalizations included in this study.

Data Example

Two examples of features sets are shown below. Figure 4 shows the normalized octave spectrum from the bird chirp example in Figure 2. All spectrum are normalized such that the power contained in the spectrum is unity and is done to remove the effects of propagation distance on the recorded levels. Darker spectrum lines indicate that the data used for the that spectrum was from later in the file. Notice that there are no distinct features.

Error! Reference source not found.5 shows the normalized octave spectrum from the aircraft example in **Error! Reference source not found.** Notice that the spectral shape is changing through the length of the time series data and that there is a distinct feature or hump around the 1 kHz octave band. For the aircraft example, the file contained 3,072,000 samples that were split into 1 second records overlapped by 25% of the record width for octave analysis resulting in 79 spectral estimates. The octave spectrum are computed for the octave bands between the 31.5 Hz band and the 16 kHz band resulting in 10 features per each spectrum. The spectrum shape from the two examples is visually different providing evidence that a machine learning algorithm should be able to separate the two feature sets as well.

The octave spectrum presented in the examples above demonstrate one of many different feature sets that could be used for classification of environmental noise. Other feature generation options include 1/n octave spectrum, which would break the octave bands seen above into smaller bands, cepstrum processing, short-time Fourier transforms, and wavelets. All the different feature sets have the potential to provide varying degrees of separation between the signals. This project should survey as many feature sets as practical to gain the understanding of which provide the best separation. Selection of the classification algorithm will also impact the system performance and should be studied as well. The various combinations of a feature set and classification algorithm will form a rich test matrix for this project to study.

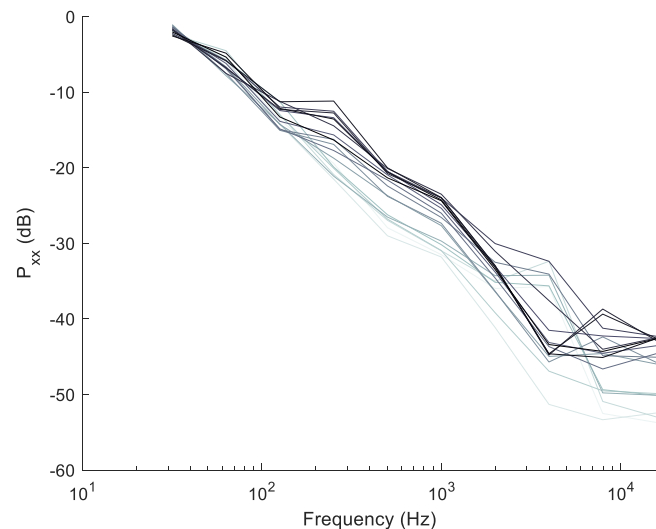


Figure 4: Octave spectrum of the bird chirp example (bird01.wav)

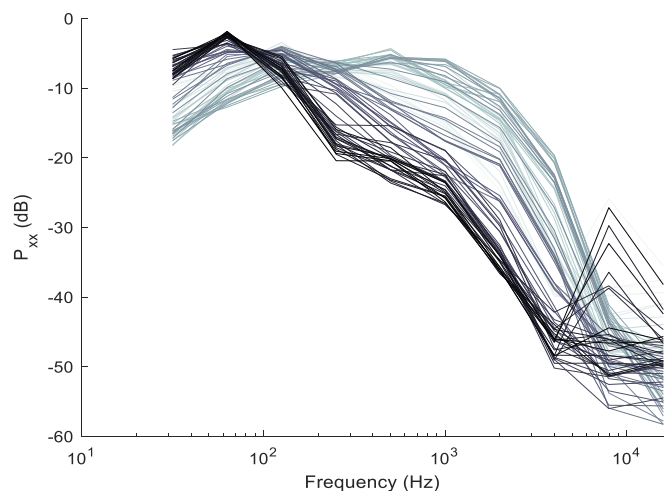


Figure 5: Octave spectrum of the aircraft example (plane08.wav).

Summary

The raw data needed for this project is represented as audio time series recordings and is readily available. The data set provided by the sponsor is representative but may need to be augmented to address the class imbalance. The data issues that are present are mostly confined to data management practices as numerous combinations of signals, processing techniques, and classifiers will need be tracked and the results reported on. With such a diverse set of features and algorithms to explore, this project presents a great learning opportunity for those interested in signal processing.

Appendix B – Project Timeline

To be able to have all our research ready to present by March 13th, 2019 (the day of the poster session), we have roughly 24 person-weeks (8 weeks X 3 people) of time to allocate to tasks for this project. Some work, such as documentation or writing a report detailing our evaluations are dependent on feature sets having been generated and models having been trained such that all of the work is not entirely parallelizable. To give ourselves extra space within this timeline, we will only propose 18 person-weeks' worth of work. The additional 6 person-weeks' worth of work will exist as a buffer for issues that arise or as time to extend the project further.

In Figure 6, we show one potential work tracking solution using Trello. Trello allows for users to manage lists of tasks and provides organizational features such as due dates, checklists for tasks and color-coded labels. Each list can be thought of as a single user story and each item on the list is a task. Add-ins for Trello allow for the addition of story points [5]. At this time, we have not agreed upon a method of applying story points to individual user stories.

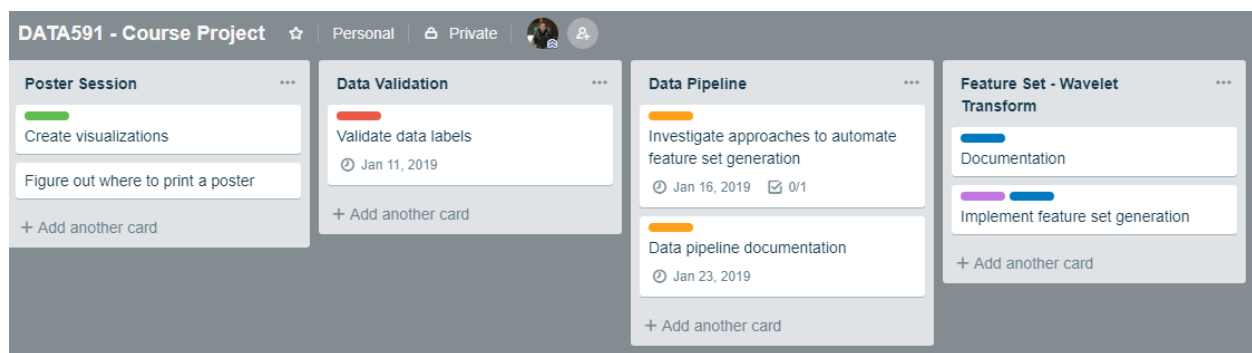


Figure 6: Example Lists on a Trello board planning work

Proposed User Stories

Data Validation/Labeling (1 person-week)

One of the initial tasks we'll have to tackle is verifying our dataset and ensuring that the labels are correct. Further annotations of the data might be necessary. For example, the labeling provided by BT&E does not include timestamp markers for when contaminating acoustic signals occur. The initial validation work has been done as part of Data Pipeline write up but further work is necessary.

Initial Data Pipeline Work (1 person-week)

Before implementing the end-to-end pipeline to manage audio files as a data source, we'll want to have some features developed that anyone on the team can leverage to perform tasks such as:

1. Ingesting audio files
2. Selecting windowed subsets of an audio file
 - a. Specifying overlap between windows
3. Maintaining appropriate data labels for the above windows

Data Pipeline (3 person-weeks)

The end-to-end data pipeline will allow a user to select a feature set and other settings they want for their covariates, point the method at a file or set of files and labels, and have the requested data set returned to be handed off to a machine learning function. This work will abstract away the need to more granularly work with feature generation.

Feature Set Generation (3 person-weeks)

For each different set of features we want to explore (Signal Filtering, Fourier Analysis, Wavelet Transformations), we'll need to perform the following tasks:

1. Learn how to generate that feature set in MATLAB
2. Document the process to generate the feature set
3. Implement generation of the feature set within our project.

While not quite an epic, this topic encompasses multiple user stories, one for each different feature set we want to explore. It is closely related to the work enumerated below for machine learning

Machine Learning Generation (7 person-weeks)

For each different machine learning technique we want to explore (Long short-term memory RNNs, softmax multiclassification, multi-layer perceptrons, logistic regression), we'll need to perform the following tasks:

1. Learn how to implement the algorithm/library in MATLAB
2. Document the process to run the algorithm
3. Train the model and evaluate performance
 - a. Do this on each feature set

This too is a collection of user stories, one for each machine learning technique. It is likely that individual machine learning techniques will have multiple members of the team working on different feature sets. This work is dependent on the prior feature set generation work.

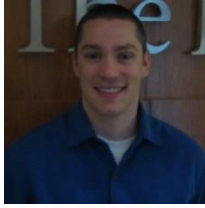
Poster Session Preparation (2 person-weeks)

Before the poster session occurs, we'll need to have generated the visualizations and context we'll want to display on our poster. This user story covers that work as well as printing the poster at a store. This story has an external dependency on whichever print shop we select and has additional time built into it to account for printing delays.

Documentation (1+ person-weeks)

This user story exists to encompass the work of writing any and all documentation/reporting that is not part of one of the above user stories. For example, the report on the comparison of different machine learning techniques and their effectiveness might fall under this story.

Appendix C – The Team



Todd Schultz

Todd brings a deep domain knowledge in acoustic signal processing and Fourier-based analysis to the team. From this project he hopes to gain practical experience with wavelet transforms and long-short term memory recurrent neural networks.



Sean Miller

Sean brings a background in data engineering to the project and is excited to operationalize the workflows we recommend to Boeing. He hopes to gain an understanding of signal processing techniques for audio that he can apply to data outside of flight tests.



Rahul Birmiwal

Rahul brings both an extensive background in signal processing/FFT algorithms and programming experience to the team. He hopes to learn about novel methods to translate acoustic signal data into an optimal set of covariates for machine-learning/classification.

Team Expectations

The team expects to sync in person once per week as part of our scheduled lecture time. For communication that occurs outside of the standup/sync, the expectation is that it will occur on Slack so all team members can contribute to discussions. We will leverage the GitHub Issue tracker to report bugs and issues that need to be addressed within our codebase.

As Todd helped bring this project to the UW Master's Degree in Data Science program, he will be responsible for reporting to and bubbling up concerns to the sponsor, Steve Underbrink

References

- [1] "Aircraft Noise Issues". *faa.gov*, 2018, https://www.faa.gov/about/office_org/headquarters_offices/apl/noise_emissions/airport_aircraft_noise_issues/. Accessed 1 Dec 2018.
- [2] "Sound Gallery - Natural Sounds (U.S. National Park Service)". *nps.gov*, 2018, <https://www.nps.gov/subjects/sound/gallery.htm>. Accessed 2 Dec 2018.
- [3] Jimena Royo-Letelier, Romain Hennequin, Viet-Anh Tran, Manuel Moussallam, "Disambiguating Music Artists at Scale with Audio Metric Learning" [arXiv:1810.01807](https://arxiv.org/abs/1810.01807) [cs.IR], Oct. 2018.
- [4] Jiyoung Park, Jongpil Lee, Jangyeon Park, Jung-Woo Ha, Juhan Nam, "Representation Learning of Music Using Artist Labels" [arXiv:1710.06648](https://arxiv.org/abs/1710.06648) [cs.SD], Oct. 2017.
- [5] "Trello - Agile Tools". *Trello.com*, <https://trello.com/power-ups/59d4ef8cfea15a55b0086614/agile-tools>. Accessed 1 Dec 2018.
- [6] Sharath Adavanne, Giambattista Parascandolo, Pasi Pertilä, Toni Heittola, Tuomas Virtanen, "Sound Event Detection in Multichannel Audio Using Spatial and Harmonic Features" [arXiv:1706.02293v1](https://arxiv.org/abs/1706.02293v1) [cs.SD], June 2017.
- [7] Y. Wang, X. Wu, X. Li, and J. Zhou. (2014) 'Vehicle type recognition in sensor networks using improved time encoding signal processing algorithm.' Mathematical Problems in Engineering, Vol 2014, Article ID 142304, 8 pgs. <http://dx.doi.org/10.1155/2014/142304>.
- [8] H. Wu, M. Siegel, P. Khosla, "Vehicle sound signature recognition by frequency vector principal component analysis," Instrumentation and Measurement Technology Conference, IEEE, 1998, pp.429-434.
- [9] Anden, J. and Mallat, S. 2014. Deep scattering spectrum. IEEE Transactions on Signal Processing, Vol. 62, 16, pp. 4114-4128.
- [10] F Hemmati, W. Orfali, M. Gadala. 2016. Roller bearing acoustic signature extraction by wavelet packet transform, applications in fault detection and size estimation. Applied Acoustic, v 104, pp. 101-118. <https://doi.org/10.1016/j.apacoust.2015.11.003>.