Research Title

Dynamic Segmentation and Feature Extraction from Non-Stationary Spatial Data Using a Hybrid Signal Decomposition and Machine Learning Techniques

By:

Abel Ayele
Instructor, Department of Mathematical Sciences
Lincoln University, PA

Total Funding Requested

Summer Salary (June & July, 2024) + Conference Travel - $7950

Application Date

02/23/2024

1. Project Description

This research aims to develop an innovative approach to address the challenges posed by non-stationary spatial data/signals. A non-stationary signal is a signal whose statistical properties vary with time. In the context of spatial signals, these characteristics change over different spatial locations or regions. This can include variations in amplitude, frequency, phase, or other properties. Temporal spatial interactions is another challenge posed by non-stationarity of signals where signal properties can change not only across different locations but also over time, introducing additional complexity. Certain features or patterns in the spatial signal may be localized to specific regions, and their characteristics may differ across the spatial domain. Identifying and characterizing these localized features pose challenges.

The goal of dynamic segmentation is to capture local variations in the signal and provide a more accurate representation for analysis. The methodology proposed in this research combines signal decomposition and machine learning techniques to enable dynamic segmentation and feature extraction, providing a comprehensive analysis of the spatial data. The hybrid approach explores advanced signal decomposition methods and integrate that with Machine Learning methods mainly clustering algorithms. Some of the signal decomposition methods to be explored in this research are:

Fourier Transform: $X(f) = \int_{-\infty}^{\infty} x(t) e^{-j2\pi f t}\, dt$

Short-Time Fourier Transform (STFT): $X(f, \tau) = \int_{-\infty}^{\infty} x(t) w(t-\tau) e^{-j2\pi f t}\, dt$

Wavelet Transform: $W(a, b) = \int_{-\infty}^{\infty} x(t) \psi^{*}\left(\frac{t-b}{a}\right)\, dt$

These equations are instrumental in decomposing the non-stationary spatial signals into their frequency components at different scales and providing a time-frequency representation. The Fourier Transform captures the frequency content of the signal in the entire time domain, while the STFT and Wavelet Transform are valuable for analyzing the time-varying frequency components of non-stationary signals. This research explores how these methods can be adopted in the spatial domain instead of time-domain. Railroad track geometry measurement data will be used in this research. Railroad track geometry refers to the physical layout and dimensions of railway tracks, including various parameters that define the alignment, curvature, and elevation.

2. Goals and Objectives of the Research

The goal of this research project is to explore advanced data analysis and machine learning methods that can be utilized to better characterize and classify non-stationary spatial data/signals using dynamic segmentation and feature extraction methods.

Objectives

1. Explore different signal decomposition techniques and implement a hybrid approach to extract intrinsic features from non-stationary spatial data.
2. Integrate machine learning algorithms for dynamic segmentation and classification of distinct segments within the signal.

3. Project Timeframe

This research project will be conducted during summer months of June and July 2024. The final report and pertinent journal/conference paper will be completed by September 31, 2024.

4. Enhancement of Teaching and Research at Lincoln University

This project will have an immense contribution to promote data science education and research at Lincoln University. The findings of this research will inform the curriculum of courses such as Statistics, Foundations of Data Science and Mathematical Statistics in the Math department. Beyond the Math department, it will help enhance the teachings of courses in other STEM fields such as Digital Signal Processing (Engineering Sciences) and Computer Programing, (such as Python). This project will also helps harness the multidisciplinary nature of data science research to foster discussion among faculty members on potential curriculum revision and research collaboration initiatives. One simple example could be the development of a joint Machine Leaning course between Math and Computer Science departments. Promoting such researches in advanced data analysis and machine learning is paramount especially in this era of Big Data and Artificial Intelligence (AI).

5. Measurement of the Project Success

The success of this project will be measured by:

1. The submission and presentation of the final report as per the outlined timeline.
2. The preparation and submission of conference paper based on the project findings

6. Sharing the Project Outcome

The project outcome will be shared through the following means:

1. The 2024/2025 faculty research symposium here at Lincoln University.
2. The annual “Big Data Conference, December 11-12, 2024, University of Delaware”.
3. Journal paper will be submitted for publication at a venue relevant to this topic.

7. Budget

Two months’ summer Salary:
June and July, 2024………$7000
Conference Travel………..$ 1950