Computers and Geosciences 114 (2018) 107–116

Contents lists available at ScienceDirect

Computers and Geosciences

journal homepage: www.elsevier.com/locate/cageo

Research paper

Parallel optimization of signal detection in active magnetospheric signal injection experiments

Michael Gowanlock [a,b,*], Justin D. Li [b], Cody M. Rude [b], Victor Pankratius [b]

[a] School of Informatics, Computing & Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA
[b] Massachusetts Institute of Technology, Haystack Observatory, Westford, MA, USA

ARTICLE INFO

Keywords:
Parallel parameter search
Signal processing
Siple station experiment
Time series analysis

ABSTRACT

Signal detection and extraction requires substantial manual parameter tuning at different stages in the processing pipeline. Time-series data depends on domain-specific signal properties, necessitating unique parameter selection for a given problem. The large potential search space makes this parameter selection process time-consuming and subject to variability. We introduce a technique to search and prune such parameter search spaces in parallel and select parameters for time series filters using breadth- and depth-first search strategies to increase the likelihood of detecting signals of interest in the field of magnetospheric physics. We focus on studying geomagnetic activity in the extremely and very low frequency ranges (ELF/VLF) using ELF/VLF transmissions from Siple Station, Antarctica, received at Québec, Canada. Our technique successfully detects amplified transmissions and achieves substantial speedup performance gains as compared to an exhaustive parameter search. We present examples where our algorithmic approach reduces the search from hundreds of seconds down to less than 1 s, with a ranked signal detection in the top 99th percentile, thus making it valuable for real-time monitoring. We also present empirical performance models quantifying the trade-off between the quality of signal recovered and the algorithm response time required for signal extraction. In the future, improved signal extraction in scenarios like the Siple experiment will enable better real-time diagnostics of conditions of the Earth's magnetosphere for monitoring space weather activity.

1. Introduction

We introduce new parallel optimization techniques for the detection of signals injected into the Earth's magnetosphere. The application addresses a variety of interesting questions, such as improving an understanding of radiation-belt dynamics, monitoring space weather conditions, and advancing scientific insight into the coupling dynamics between the Earth and the Sun.

Signal detection and extraction is challenging for applications where signal transmission occurs in noisy channels with interference. This is further complicated in this application by interactions of interest between energetic particles and whistler-mode waves in the radiation belts that modify wave behavior and particle populations (Harid et al., 2014a, 2014b; Streltsov et al., 2010; Li et al., 2015). These waves are generated on Earth naturally, for instance by lightning, or artificially, such as by power lines. We examine data from a controlled wave-particle interaction experiment at Siple Station, Antarctica, where a dedicated transmitter injected signals in the extremely and very low frequency ranges (ELF/VLF; 0.3–30 kHz) into the magnetosphere. Such signals travel along field-aligned paths through the magnetosphere, undergoing modifications through the interactions with energetic electrons in the radiation belts, and are received at the geomagnetic conjugate point in Québec, Canada (Li et al., 2014; Helliwell and Katsufrakis, 1974). This experimental setup is shown in Fig. 1.

The received signal is broadband in nature, which requires applying a configurable signal processing chain in order to extract the signal of interest, including for example a lightning filter, a demodulation stage, a smoothing filter, and a signal thresholding filter. The challenge is: how do we select the respective parameters in each stage in order to separate the noise and enhance the detection of the transmitted signal? This problem can lead to large search spaces which can make an exhaustive search intractable.

While the typical signal detection approach uses a manually tuned signal processing chain, alternate parameters from the broader permissible range could provide other insights and emphasize different salient features in the signal. In this work, we automate this search and pruning

* Corresponding author. School of Informatics, Computing & Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA.
E-mail addresses: michael.gowanlock@nau.edu (M. Gowanlock), jdli@mit.edu (J.D. Li), cmrude@mit.edu (C.M. Rude), pankrat@mit.edu (V. Pankratius).

https://doi.org/10.1016/j.cageo.2018.01.020
Received 10 March 2017; Received in revised form 7 August 2017; Accepted 30 January 2018
Available online 11 February 2018

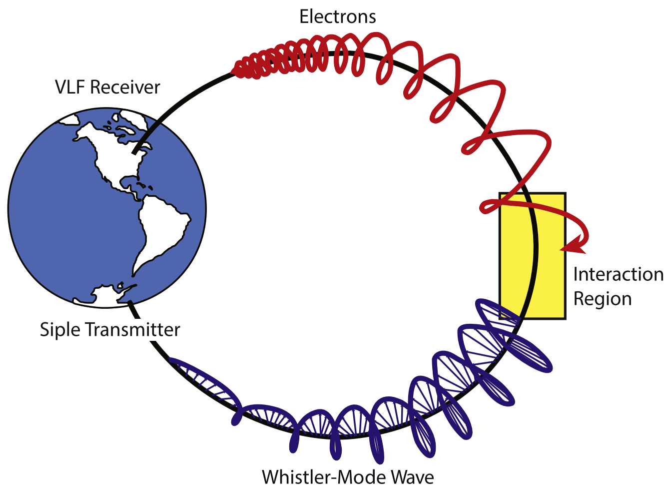

Fig. 1. The transmitter at Siple Station, Antarctica, injects whistler-mode waves (blue) into the magnetosphere, which undergo interactions with energetic electrons (red) in the interaction region (yellow) near the equator. The modified waves then continue along the field-aligned ducts and are received at the receiver at the geomagnetic conjugate point in Québec, Canada. This figure is adapted from (Li, 2015). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

process while leveraging parallel computing to improve algorithm response time. These performance gains can lead to better scientific insights by exploring more of the parameter space, or enable real-time signal detection. We develop a parallel, multithreaded implementation that achieves good scalability on a multicore platform which improves search throughput and present the results from our evaluation on data from the Siple Station Experiment. Our methods show possible applications for improving the detection of events exhibiting magnetospheric amplification and for highlighting different salient features based on the searched parameter space.

2. The Siple station magnetospheric signal injection experiment

2.1. Overview

Siple Station, Antarctica, operated a powerful ELF/VLF transmitter for controlled wave-injection experiments from 1973 to 1988. It enabled a unique experiment in magnetospheric physics, whose observations remain unmatched by any other instrument in providing long-running observations of wave-particle interactions in the magnetosphere (Li, 2015). These data have improved our understanding of various magnetospheric phenomena and processes, but due to their historicity and format, remain an underutilized dataset (Li et al., 2014).

We examine data from the 1986 portion of the dataset, which was previously preprocessed to allow for digital analysis (Li et al., 2014). An example of an MDIAG (magnetospheric diagnostic format) transmission from 6/23/1986 7:01:00 UT is shown in spectrogram format in Fig. 2. The spectrogram shows several features of interest that must be considered in the signal processing analysis. The narrow, vertical features indicate lightning strikes, while the long, horizontal features correspond with power line harmonics. Both such features interfere with detecting the transmitted signal. The 2 s pulse of interest is the higher amplitude signal at 3 480 Hz, which can be seen starting around 3 s.

2.2. Siple experiment signal processing pipeline

The canonical signal analysis in this field is exemplified by the process in (Li et al., 2014; Li et al., 2015). It consists of removing the lightning-generated noise (called sferics) in the data, extracting the narrowband signal amplitude based on transmission characteristics,

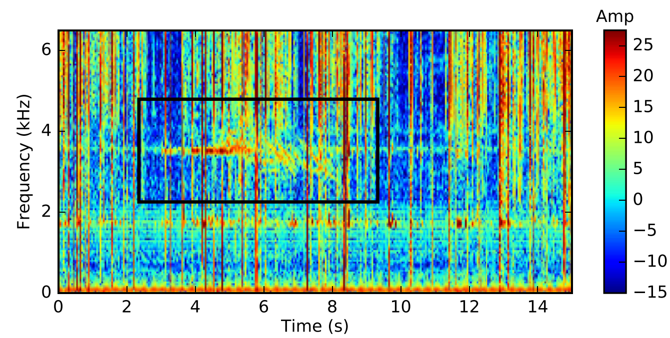

Fig. 2. Example of an MDIAG transmission from 6/23/1986 7:01:00 UT shown in a time-frequency amplitude representation using a spectrogram. The signal component is the narrowband transmission at 3 480 Hz, as part of the magnetospheric diagnostic format, MDIAG (Li et al., 2014) which is enclosed within the black box.

smoothing to reduce noise, and thresholding to generate a simplified signal for computing the fitness function. Sferics are strongly impulsive and can easily dominate the received broadband signal. For simplicity, sferics are primarily removed by zeroing portions of the signal that exceed some number of standard deviations of the amplitude. Then, the signal is mixed down to baseband by shifting the given transmission frequency to 0 Hz and low-pass filtering to extract the narrowband amplitude. A median filter is applied to smooth residual impulsive noise from the sferics as well as artifacts from the low-pass filter. Finally, a regression tree classifier (Breiman et al., 1984) thresholds the signal and approximates the detectable signal amplitudes.

The quality of the signal extraction is evaluated by cross-correlating the thresholded signal with a defined approximation of the signal, which can be obtained based on the transmission characteristics. The result is normalized to range between −1 and 2, which will serve as a fitness function score in our approach (Section 2.2.3).

2.2.1. Signal processing pipeline stages

- Stage 1: Takes the entire dataset and thresholds by a single parameter, the number of standard deviations, to remove the impulsive sferics and produce a filtered dataset of the same length. The number of standard deviations ranges between 0.5 and 2.0 to encompass the range of coincident sferics (Said, 2009).
- Stage 2: Takes the sferic-filtered data and calculates the narrowband amplitude at the known transmission frequency. This stage involves 4 steps: (a) mixing the signal down to baseband using a complex exponential frequency shift; (b) generating a low-pass filter using three parameters, the number of filter taps, the passband frequency, and the roll-off frequency; (c) applying the low-pass filter; and (d) downsampling the data based on passband frequency. The output of this stage is the downsampled narrowband signal amplitude. Only the low-pass filter, defined using the Remez algorithm (McClellan and Parks, 1973), requires parameters. The number of taps ranges from 100 to 300, and the cut-off and roll-off frequencies each range from 100 to 200 Hz. These parameter ranges are reasonable for the Siple signal processing chain, similar to the values used in (Li et al., 2014).
- Stage 3: Takes the downsampled data and applies a naïve median filter (Robinson, 2004) with a parameter-specified window size to smooth the narrowband amplitude. The downsampling typically shrinks the data by 10 − 100x. The output of this stage is a smoothed amplitude, where the data length is unchanged. Window sizes range from 7 to 23, encompassing the value used in (Li et al., 2014).
- Stage 4: Fits a regression tree (Breiman et al., 1984) of depth specified by the tree-depth parameter to estimate the time characteristics of the signal. The regression tree output of this stage is used to compute the fitness function score that determines the quality of the parameters