Title: Anomaly Detection in Earth Dam and Levee Passive Seismic Data Using Support Vector Machines and Automatic Feature Selection

Author: Wendy D. Fisher Tracy K. Camp Valeria V. Krzhizhanovskaya

PII: S1877-7503(16)30418-5
DOI: http://dx.doi.org/doi:10.1016/j.jocs.2016.11.016
Reference: JOCS 586

To appear in:

Received date: 15-9-2016
Revised date: 18-11-2016
Accepted date: 27-11-2016

Please cite this article as: Wendy D. Fisher, Tracy K. Camp, Valeria V. Krzhizhanovskaya, Anomaly Detection in Earth Dam and Levee Passive Seismic Data Using Support Vector Machines and Automatic Feature Selection, *<![CDATA[Journal of Computational Science]]>* (2016), http://dx.doi.org/10.1016/j.jocs.2016.11.016

This space is reserved for the Procedia header, do not use it

Anomaly Detection in Earth Dam and Levee Passive Seismic Data Using Support Vector Machines and Automatic Feature Selection

Wendy D. Fisher[1], Tracy K. Camp[1], and Valeria V. Krzhizhanovskaya[2,3,4]

[1] Colorado School of Mines, Golden, Colorado, U.S.A.
wbelcher@mines.edu, tcamp@mines.edu
[2] University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, the Netherlands
V.Krzhizhanovskaya@uva.nl
[3] National Research University ITMO, St. Petersburg, 197101, Russia
[4] St. Petersburg State Polytechnic University, St. Petersburg, 195251, Russia

Abstract

We investigate techniques for earth dam and levee health monitoring and automatic detection of anomalous events in passive seismic data. We have developed a novel data-driven workflow specific to our domain, which could be generalized for monitoring other systems with time series data. We use machine learning and geophysical data collected from sensors located on the surface of the levee to identify internal erosion events. In this paper, we describe our research experiments with two-class and one-class Support Vector Machines (SVMs). We use two different data sets from experimental laboratory earth embankments (each having approximately 80% normal and 20% anomalies) to ensure our workflow is robust enough to work with multiple data sets and different types of anomalous events (e.g., cracks and piping). We apply wavelet-denoising techniques and extract nine spectral features from decomposed segments of the time series data. The two-class SVM with 10-fold cross validation achieved over 94% overall accuracy and 96% F1-score. Experiments with the one-class SVM (no labeled data for anomalies) using the top features selected by our automatic feature selection algorithm increase our overall results from 83% accuracy and 89% F1-score to over 91% accuracy and 95% F1-score. Results show that we can successfully separate normal from anomalous data observations.

This is an extended version of our conference paper that was invited to the JoCS special issue (http://dx.doi.org/10.1016/j.procs.2016.05.339).

Keywords: Data-driven levee monitoring, machine learning, anomaly detection, passive seismic.