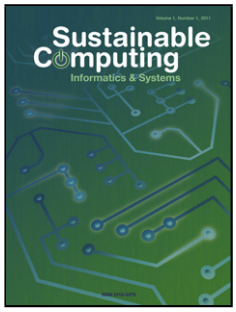

Title: Exploring the Capabilities of Support Vector Machines in Detecting Silent Data Corruptions

Author: Omer Subasi Sheng Di Leonardo Bautista-Gomez Prasanna Balaprakash Osman Unsal Jesus Labarta Adrian Cristal Sriram Krishnamoorthy Franck Cappello

PII: S2210-5379(17)30089-6
DOI: https://doi.org/doi:10.1016/j.suscom.2018.01.004
Reference: SUSCOM 223

To appear in:

Received date: 17-3-2017
Revised date: 4-10-2017
Accepted date: 18-1-2018

Please cite this article as: Omer Subasi, Sheng Di, Leonardo Bautista-Gomez, Prasanna Balaprakash, Osman Unsal, Jesus Labarta, Adrian Cristal, Sriram Krishnamoorthy, Franck Cappello, Exploring the Capabilities of Support Vector Machines in Detecting Silent Data Corruptions, *<![CDATA[Sustainable Computing: Informatics and Systems]]>* (2018), https://doi.org/10.1016/j.suscom.2018.01.004

Exploring the Capabilities of Support Vector Machines in Detecting Silent Data Corruptions

Omer Subasi,[1] Sheng Di,[3] Leonardo Bautista-Gomez,[2] Prasanna Balaprakash,[3]
Osman Unsal,[2] Jesus Labarta,[2] Adrian Cristal,[2,4] Sriram Krishnamoorthy,[1]
Franck Cappello[3]
[1]Pacific Northwest National Laboratory, Washington, USA
[2]Barcelona Supercomputing Center, Spain
[3]Argonne National Laboratory, Lemont, Illinois, USA
[4]IIIA - Artificial Intelligence Research Institute CSIC - Spanish National Research Council, Spain
{omer.subasi, sriram}@pnnl.gov
{leonardo.bautista, osman.unsal, jesus.labarta, adrian.cristal}@bsc.es
{sdi1, cappello, pbalapra}@anl.gov

Abstract

As the exascale era approaches, the increasing capacity of high-performance computing (HPC) systems with targeted power and energy budget goals introduces significant challenges in reliability. Silent data corruptions (SDCs), or silent errors, are one of the major sources that corrupt the execution results of HPC applications without being detected.

In this work, we explore a set of novel SDC detectors - by leveraging epsilon-insensitive support vector machine regression - to detect SDCs that occur in HPC applications. The key contributions are threefold. (1) Our exploration takes temporal, spatial, and spatiotemporal features into account and analyzes different detectors based on different features. (2) We provide an in-depth study on the detection ability and performance with different parameters, and we optimize the detection range carefully. (3) Experiments with eight real-world HPC applications show that support-vector-machine-based detectors can achieve detection sensitivity (i.e., recall) up to 99% yet suffer a less than 1% false positive rate for most cases. Our detectors incur low performance overhead, 5% on average, for all benchmarks studied in this work.

Keywords: Silent Data Corruptions, Support Vector Machines, HPC Applications.