NSF DIBBs PI Workshop, Jan 11-12, 2017

DIBBS ACI-1442997

An Infrastructure for Computer Aided Discovery in Geoscience

Victor Pankratius[1] (PI), Philp J. Erickson (coPI), Frank D. Lind (coPI),
Michael Gowanlock, Cody M. Rude, Justin D. Li, Guillaume Rongier

Massachusetts Institute of Technology, Haystack Observatory

[1]Contact: pankrat@mit.edu

Abstract

Next-generation Geoscience needs to handle rapidly growing data volumes from ground-based and space-based sensor networks. As real-world phenomena are mapped to data, the scientific discovery process essentially becomes a search process across multidimensional data sets. The extraction of meaningful discoveries from this sea of data therefore requires scalable machine assistance to enhance human contextual understanding.

This project develops a computer-aided discovery methodology and infrastructure that provides scientists with better support for scientific question answering. The pragmatics of our model-based discovery system go beyond feature detection in empirical data to answer fundamental questions, such as how empirical detections fit into hypothesized models and model variants to ease the scientist's work of placing large ensembles of detections into a theoretical context. To achieve this, scientists can programmatically express hypothesized scenarios, constraints, and model variations in a cloud environment. This approach helps delegate the automatic exploration of the combinatorial search space of possible explanations in parallel on a variety of data sets.

We demonstrate successful applications of this paradigm in several areas such as space weather and ionospheric studies, volcanics and surface deformation, and discuss further generalizations of our approach for other science areas.

Selected References

- ***Computer-Aided Discovery: Towards Scientific Insight Generation with Machine Support.*** *Pankratius, V., Li, J., Gowanlock, M., Blair, D. M., Rude, C., Herring, T., Lind, F., Erickson, J., & Lonsdale, C. IEEE Intelligent Systems, 31(4): 3–10. 2016*
  http://doi.ieeecomputersociety.org/10.1109/MIS.2016.60

- ***Exploiting Variant-Based Parallelism for Data Mining of Space Weather Phenomena****. Gowanlock, M., Blair, D. M. & Pankratius, V., 30th IEEE International Parallel & Distributed Symposium (IPDPS 2016) http://dx.doi.org/10.1109/IPDPS.2016.10*

- ***Computer Aided Detection of Transient Inflation Events at Alaskan Volcanoes using GPS Measurements from 2005-2015.*** *Li, J. D., Rude, C. M., Blair, D. M., Gowanlock, M. G., Herring, T. A. & Pankratius, V., Journal of Volcanology and Geothermal Research, accepted Oct 4, 2016 – in press.*
  http://dx.doi.org/10.1016/j.jvolgeores.2016.10.003

- ***Machine Learning in Ionospheric Phenomena Detection Using Passive Radar.*** *Barrari, S., Pankratius, V., Lind, F. 46th Annual Fall Meeting, ID 65374, Session IN51A-1781 (Big Data Analytics for Science Data I), Fri Dec 18, 2015*
  https://agu.confex.com/agu/fm15/mediafile/Handout/Paper65374/IN51A-1781.pdf