Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by **Battelle** *Since 1965*

Science

Advanced Computing, Mathematics and Data Research Highlights

August 2017

Data, HPC Scientists Team for GraphChallenge Championship
Grappolo graph clustering library impresses with code quality

A team featuring **Mahantesh Halappanavar** (Data Sciences) and **Antonino Tumeo** (High Performance Computing), both from PNNL's Advanced Computing, Mathematics, and Data Division, along with collaborators Hao Lu (Oak Ridge National Laboratory) and Ananth Kalyanaraman (Washington State University), was named GraphChallenge champions for their work "Scalable Static and Dynamic Community Detection Using Grappolo."

The GraphChallenge was created to encourage development of "new solutions for analyzing graphs derived from social media, sensor feeds, and scientific data to enable relationships between events to be discovered as they unfold in the field." Such innovations in graph analytics will support efforts such as the Defense Advanced Research Projects Agency (DARPA)'s Hierarchical Identify Verify Exploit, or HIVE, program. HIVE's goal is to develop a single graph analytics processor that can achieve 1,000 times improvement in data processing efficiency. This capability is expected to have broad applications to varied problems in areas such as cybersecurity, social media analysis, and infrastructure modeling.

Mahantesh Halappanavar (left) and Antonino Tumeo (right) were part of a team named GraphChallenge champions. Enlarge Image.

In recognition of their achievement, the team will present a 10-minute talk about their work at the 2017 IEEE High Performance Extreme Computing Conference (HPEC) being held September 12-14, 2017 in Waltham, Massachusetts. Their work also will appear in the official IEEE HPEC conference proceedings.

According to Halappanavar, a primary developer of Grappolo, a multithreaded C++ and OpenMP library for graph clustering (community detection) based on the Louvain method, the entire team was pleased by the recognition but even more so by the quality of the code.

"During the submission, we were also pleasantly surprised by the quality of output Grappolo achieved," he explained. "It was almost 100 percent in Precision and Recall for most of the challenge inputs with ground truth data."

Enlarge Image.

Currently, Grappolo is being tested on real-world data sets from biology.

Acknowledgments