Guest Editorial: Big Data Analytics and the Web

Michael Sheng, Athanasios V. Vasilakos, Qi Yu, and Lina You

The paper by Shao et al., “Clustering Big Spatiotemporal-Interval Data,” focuses on clustering big spatiotemporal data, which are common in the emerging Web of Things (WoT), where a large number of sensors are deployed for continuously collecting data. The authors explore a novel way to cluster massive Web data with spatiotemporal intervals in multiple Euclidean spaces, as well as a new energy function to measure similarity and balance between clusters in spatial, temporal, and data dimensions.

The paper by Zhao et al., “A Distributed Graph-Parallel Computing System with Lightweight Communication Overhead,” targets complex and large-scale graph processing and develops a system called Ligraph that provides three new features, namely a Gather partial sum difference based computing model, a corresponding lightweight Gather communication mechanism, and a lightweight synchronizing communication mechanism and an edge direction-aware graph partition strategy for PageRank-like algorithms.

The paper by Wang et al., “Kvasir: Scalable Provision of Semantically Relevant Web Content on Big Data Framework,” targets effective content-based filtering which is critical to today’s Internet where users are overloaded with excessive information flows. Smart content provision and recommendation become more and more crucial in improving user experience and efficiency in using Internet applications. The authors present a scalable semantic recommendation system named Kvasir to seamlessly integrate an automated and proactive content provision service into web browsing.

The paper by Chen, Kazman, and Haziyev, “Agile Big Data Analytics for Web-based Systems: An Architecture-Centric Approach”, introduces an architecture-centric approach, called Architecture-centric Agile Big data Analytics (AABA) to address technical and organizational challenges in big data system development and agile delivery of big data analytics for Web-based systems. Eleven case studies have been discussed in this article, which demonstrate that architecture-centric design, development, and operation is the key to taming technical complexity, as well as achieving necessary agility for big data analytics for Web-based systems.

For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below. Digital Object Identifier no. 10.1109/TBDATA.2016.2613758