World Academy of Science, Engineering and Technology
International Journal of Health and Medical Engineering
Vol:10, No:05, 2016

Indexing and Incremental Approach Using Map Reduce Bipartite Graph (MRBG) for Mining Evolving Big Data

Authors : Adarsh Shroff

Abstract : Big data is a collection of dataset so large and complex that it becomes difficult to process using data base management tools. To perform operations like search, analysis, visualization on big data by using data mining; which is the process of extraction of patterns or knowledge from large data set. In recent years, the data mining applications become stale and obsolete over time. Incremental processing is a promising approach to refreshing mining results. It utilizes previously saved states to avoid the expense of re-computation from scratch. This project uses i2MapReduce, an incremental processing extension to Map Reduce, the most widely used framework for mining big data. I2MapReduce performs key-value pair level incremental processing rather than task level re-computation, supports not only one-step computation but also more sophisticated iterative computation, which is widely used in data mining applications, and incorporates a set of novel techniques to reduce I/O overhead for accessing preserved fine-grain computation states. To optimize the mining results, evaluate i2MapReduce using a one-step algorithm and three iterative algorithms with diverse computation characteristics for efficient mining.

Keywords : big data, map reduce, incremental processing, iterative computation
Conference Title : ICHQM 2016 : International Conference on Healthcare Quality Management
Conference Location : London, United Kingdom
Conference Dates : May 23-24, 2016