Abstract

The paper presents a new Cyber Physical Stream algorithm for selecting a predominant item from very large collections of data. The algorithm effectively works for frequencies of the predominant items starting from about 2%. The algorithm is focused on querying massive data in Software-Defined Storage combined with Fuzzy indexing method. Experiment results show that Cyber Physical Stream algorithm improves the accuracy and efficiency over previous efforts.

"IBM software defined storage", http://www-03.ibm.com/systems/storage/software-defined-storage


Index Terms

Computer Science

Algorithms

Keywords

Big Data processing; Stream Algorithms; Software-Defined Storage; Majority algorithm; Fuzzy search; Database Management Systems; Pigeonhole principle.