

Accepted Manuscript

Bloom filter based optimization scheme for massive data handling in IoT environment

Amritpal Singh, Sahil Garg, Shalini Batra, Neeraj Kumar, Joel J.P.C. Rodrigues



PII: S0167-739X(17)31451-6
DOI: <https://doi.org/10.1016/j.future.2017.12.016>
Reference: FUTURE 3854

To appear in: *Future Generation Computer Systems*

Received date : 30 June 2017
Revised date : 22 September 2017
Accepted date : 14 December 2017

Please cite this article as: A. Singh, S. Garg, S. Batra, N. Kumar, J.J.P.C. Rodrigues, Bloom filter based optimization scheme for massive data handling in IoT environment, *Future Generation Computer Systems* (2017), <https://doi.org/10.1016/j.future.2017.12.016>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Bloom Filter based Optimization Scheme for Massive Data Handling in IoT Environment

Amritpal Singh^a, Sahil Garg^a, Shalini Batra^a, Neeraj Kumar^a, Joel J.P.C. Rodrigues^b

^aComputer Science & Engineering Department, Thapar Institute of Engineering and Technology (Deemed University), Patiala (Punjab), India.

^bNational Institute of Telecommunications (Inatel); Instituto de Telecomunicaes, Universidade da Beira Interior, Portugal; University ITMO, Russia; and University of Fortaleza (UNIFOR), Brazil.

Abstract

With the widespread popularity of big data usage across various applications, need for efficient storage, processing, and retrieval of massive data sets generated from different applications has become inevitable. Further, handling of these datasets has become one of the biggest challenges for the research community due to the involved heterogeneity in their formats. This can be attributed to their diverse sources of generation ranging from sensors to on-line transactions data and social media access. In this direction, probabilistic data structures (PDS) are suitable for large-scale data processing, approximate predictions, fast retrieval and unstructured data storage. In conventional databases, entire data needs to be stored in memory for efficient processing, but applications involving real time in-stream data demand time-bound query output in a single pass. Hence, this paper proposes Accommodative Bloom filter (ABF), a variant of scalable bloom filter, where insertion of bulk data is done using the addition of new filters vertically. Array of m bits is divided into b buckets of l bits each and new filters of size ' m/k ' are added to each bucket to accommodate the incoming data. Data generated from various sensors has been considered for experimental purposes where query processing is done at two levels to improve

Email addresses: amritpal.singh203@gmail.com (Amritpal Singh), garg.sahil1990@gmail.com (Sahil Garg), sbatra@thapar.edu (Shalini Batra), neeraj.kumar@thapar.edu (Neeraj Kumar), joeljr@ieee.org (Joel J.P.C. Rodrigues)

Download English Version:

<https://daneshyari.com/en/article/6873219>

Download Persian Version:

<https://daneshyari.com/article/6873219>

[Daneshyari.com](https://daneshyari.com)