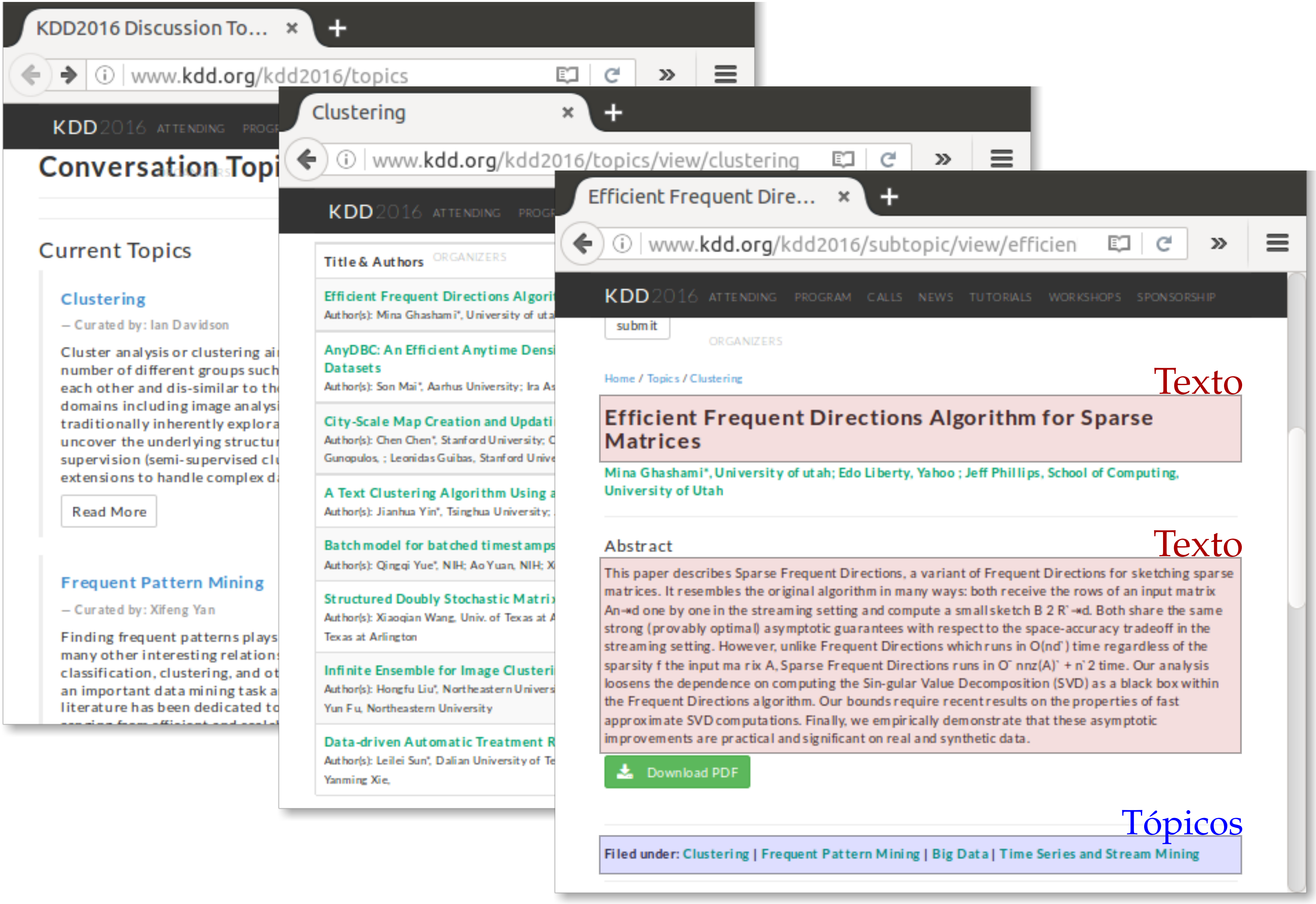
KDD2016 Discussion To...
www.kdd.org/kdd2016/topics
Conversation Topi
Current Topics
Clustering
— Curated by: Ian Davidson
Read More
Frequent Pattern Mining
— Curated by: Xifeng Yan
Clustering
www.kdd.org/kdd2016/topics/view/clustering
Title & Authors
Efficient Frequent Directions Algori
AnyDBC: An Efficient Anytime Densi
City-Scale Map Creation and Updati
A Text Clustering Algorithm Using a
Batch model for batched timestamps
Structured Doubly Stochastic Matrix
Infinite Ensemble for Image Clusteri
Data-driven Automatic Treatment R
Efficient Frequent Dire...
www.kdd.org/kdd2016/subtopic/view/efficien
KDD2016 ATTENDING PROGRAM CALLS NEWS TUTORIALS WORKSHOPS SPONSORSHIP
submit
ORGANIZERS
Home / Topics / Clustering
Texto
Efficient Frequent Directions Algorithm for Sparse Matrices
Mina Ghashami*, University of utah; Edo Liberty, Yahoo ; Jeff Phillips, School of Computing, University of Utah
Abstract
Texto
This paper describes Sparse Frequent Directions, a variant of Frequent Directions for sketching sparse matrices. It resembles the original algorithm in many ways: both receive the rows of an input matrix one by one in the streaming setting and compute a small sketch. Both share the same strong (provably optimal) asymptotic guarantees with respect to the space-accuracy tradeoff in the streaming setting. However, unlike Frequent Directions which runs in O(nd) time regardless of the sparsity f the input ma rix A, Sparse Frequent Directions runs in time. Our analysis loosens the dependence on computing the Sin-gular Value Decomposition (SVD) as a black box within the Frequent Directions algorithm. Our bounds require recent results on the properties of fast approximate SVD computations. Finally, we empirically demonstrate that these asymptotic improvements are practical and significant on real and synthetic data.
Download PDF
Tópicos
Filed under: Clustering | Frequent Pattern Mining | Big Data | Time Series and Stream Mining