

Examples for Project Topics

- 1) Out-of-core data structures for priority queues, dictionaries, hashing, etc.
- 2) Out-of-core algorithms for the construction of suffix arrays and suffix trees
- 3) Out-of-core algorithms for graph problems
- 4) Out-of-core algorithms for clustering
- 5) Out-of-core algorithms for sequence analysis
- 6) Biological data compression (DNA and protein sequences, for example)
- 7) Biological data reduction techniques
- 8) Out-of-core algorithms for rules mining
- 9) Out-of-core algorithms for matrix problems
- 10) Out-of-core algorithms for motif search (planted motif search, for example)
- 11) Efficient algorithms for k-mer counting
- 12) Efficient algorithms for sequence assembly
- 13) Efficient algorithms for the closest pair problem
- 14) Mansurul Bhuiyan and Mohammad Al Hasan, PRIIME: A Generic Framework for Interactive Personalized Interesting Pattern Discovery, IEEE Big Data 2016.
- 15) Mehrdad Yazdani, Bryn Taylor, Justine Debelius, Weizhong Li, Rob Knight, and Larry Smarr, Using Machine Learning to Identify Major Shifts in Human Gut Microbiome Protein Family Abundance in Disease, IEEE Big Data 2016.
- 16) Xiaoyi Lu, Dipti Shankar, Shashank Gugnani, and Dhableswar K. Panda, High-Performance Design of Apache Spark with RDMA and Its Benefits on Various Workloads, IEEE Big Data 2016.
- 17) Shiblee Sadik, Le Gruenwald, and Eleazar Leal, In Pursuit of Outliers in Multi-dimensional Data Streams, IEEE Big Data 2016.
- 18) Sarasi Lalithsena, Pavan Kapanipathi, and Amit Sheth, Harnessing Relationships for Domain-specific Subgraph Extraction: A Recommendation Use Case, IEEE Big Data 2016.
- 19) Chunqiu Zeng, Qing Wang, Wentao Wang, Tao Li, and Larisa Shwartz, Online Inference for Time-varying Temporal Dependency Discovery from Time Series, IEEE Big Data 2016.
- 20) Azad Naik and Huzefa Rangwala, Embedding Feature Selection for Large-scale Hierarchical Classification, IEEE Big Data 2016.
- 21) Morteza Zihayat, Zane Zhenhua Hu, Aijun An, and Yonggang Hu, Distributed and Parallel High Utility Sequential Pattern Mining, IEEE Big Data 2016.
- 22) . Xiaodong Yu, Kaixi Hou, Hao Wang, and Wu-chun Feng, Hierarchical Automata Construction for Approximate Pattern Matching on Automata Processors, IEEE Big Data 2017.
- 23) . Ichitaro Yamazaki, Stanimire Tomov, and Jack Dongarra, Sampling Algorithms to Update Truncated SVD, IEEE Big Data 2017.
- 24) Tong Yang, Binchao Yin, Hang Li, Muhammad Shahzad, Steve Uhlig, Bin Cui, and Xiaoming Li, Rectangular Hash Table: Bloom Filter and Bitmap Assisted Hash Table with High Speed, IEEE Big Data 2017.
- 25) Yueyao Wang, Qinmin Vivian Hu, Yang Song, and Liang He, Potentiality of Healthcare Big data: Improving Search by Automatic Query Reformulation, IEEE Big Data 2017.

- 26) Natalia Ponomareva, Thomas Colthurst, Gilbert Hendry, Salem Haykal, and Soroush Radpour, Compact Multi-Class Boosted Trees, IEEE Big Data 2017.
- 27) . Arnab K. Paul, Arpit Goyal, Feiyi Wang, Sarp Oral, Ali R. Butt, Michael J. Brim, and Sangeetha B. Srinivasa, I/O Load Balancing for Big Data HPC Applications, IEEE Big Data 2017.
- 28) Shashank Gugnani, Xiaoyi Lu, Houliang Qi, Li Zha, and Dhabaleswar K. Panda, Characterizing and Accelerating Indexing Techniques on Distributed Ordered Tables, IEEE Big Data 2017.
- 29) Philipp Baumann, Dorit Hochbaum, and Quico Spaen, High-Performance Geometric Algorithms for Sparse Computation in Big Data Analytics, IEEE Big Data 2017.
- 30) . Yizhou Yan, Lei Cao, and Elke Rundensteiner, Distributed Top-N Local Outlier Detection in Big Data, IEEE Big Data 2017.
- 31) Chuxu Zhang, Lu Yu, Xiangliang Zhang, and Nitesh Chawla, ImWalkMF: Joint Matrix Factorization and Implicit Walk Integrative Learning for Recommendation, IEEE Big Data 2017.
- 32) Haekyu Park, Jinhong Jung, and U Kang, A Comparative Study of Matrix Factorization and Random Walk with Restart in Recommender Systems, IEEE Big Data 2017.
- 33) Lorenzo De Stefani, Erisa Terolli, and Eli Upfal, Tiered Sampling: An Efficient Method for Approximate Counting Sparse Motifs in Massive Graph Streams, IEEE Big Data 2017.
- 34) Ankit Desai and Sanjay Chaudhary, Distributed Decision Tree v.2.0, IEEE Big Data 2017.
- 35) Byron Gao, Robert Tung, and Yong Yang, Iterative Matrix Correlation for Bisection Clustering, IEEE Big Data 2017.
- 36) Xinhui Tian, Yuanqing Guo, and Jianfeng Zhan, Towards Memory and Computation Efficient Graph Processing on Spark, IEEE Big Data 2017.
- 37) Celestine DÄanner, Thomas Parnell, Kubilay Atasü, Manolis Sifalakis, and Haralampos Pozidis, Understanding and Optimizing the Performance of Distributed Machine Learning: Applications on Apache Spark, IEEE Big Data 2017.
- 38) Lars Arge, Mathias Rav, Svend C. Svendsen, and Jakob Truelsen, External Memory Pipelining Made Easy With TPIE, IEEE Big Data 2017.
- 39) Hung Tran-The and Koji Zettsu, Discovering Co-occurrence Patterns of Heterogeneous Events from Unevenly-distributed Spatiotemporal Data, IEEE Big Data 2017.
- 40) Alexander Denzler and Michael Kaufmann, Toward Granular Knowledge Analytics for Data Intelligence, IEEE Big Data 2017.