Ablimit Aji

http://ablimit.github.io

Research Summary

My primary expertise is in research, design, and development of *large scale data management systems* and *analytics infrastructure* that spans parallel database systems and MapReduce based distributed systems. I also have extensive skills in information retrieval, data mining and machine learning.

Education

Ph.D. in Computer Science, **Emory University**, Atlanta, GA, USA   2008-2014

   Dissertation: High Performance Query Processing for Large Scale Spatial Data Warehousing
   Adviser: Prof. Fusheng Wang

M.S. in Computer Science, **Emory University**, Atlanta, GA, USA   2008-2010

   Project Topic: Information Retrieval in the Context of Social Media

B.E. in Computer Science (summa cum laude), **University of Science & Technology Beijing**   2002-2006

   Thesis title: Exploration of Machine Learning Techniques for Traditional Chinese Medicine

Professional Experience

Research Scientist – HP Labs, Palo Alto, USA   April 2014 - now

Research Assistant – Emory University, Atlanta, USA   July 2008 - March 2014

   ◇ I mainly *conceived*, *designed* and *implemented* a MapReduce based spatial query system – HadoopGIS.
   ◇ Open sourced the system as *HiveSP* – a Hive based release, and an extensible library release *libhadoopgis*.
   ◇ Published several reserach papers about the framework and system, and the work is adopted by an industrial collaborator.

Engineering Intern – Yahoo!, Sunnyvale, USA   May 2011 - Jul 2011

   During my internship (followed by a job offer), I extended the existing Infinit Browse news recommendation system to utilize named entities mentioned in the news context to improve the ranking algorithm. The code I wrote is integrated into the Infinite Browse codebase.

Research Intern – Microsoft Research, Redmond, USA   Jun 2010 - Aug 2010

   Proposed and implemented a generative modeling approach to detect and disambiguate the people named entities in web pages, and tag those entities for improving search engine relevancy. A patent for the work is filed by Microsoft.

Research Intern – Max-Planck Institute for Informatics, Saarbrücken, Germany   Jun 2009 - Sep 2009

   Redesigned the *TopX 2.0* XML query engine to perform indexing and searching tasks in *parallel*. After the redesign, the system achieved huge efficiency and ranked among the top 3 system that participating the competition.

Engineering Intern – Microsoft, Beijing, China   Oct 2007 - Feb 2008

   I co-developed an add-in for Microsoft SharePoint Server 2007, which used for an internal prototype.

Research Engineer – USTB Data Mining Lab, Beijing, China   Dec 2005 - Oct 2007

   Led a team of 3 graduate students to develop two web applications.

Publications

Journals & Book Chapters

◇ **A. Aji**, H. Vo, F. Wang: Spatial Queries in the Cloud. (Encyclopedia of Database Systems, 2015)

◇ F. Wang, **A. Aji**, G. Theodoro: Medical Image Dataset Processing over Cloud/MapReduce with Heterogeneous Architectures. (Encyclopedia of GIS, 2015)

Referred Conference and Workshop Papers

◇ H. Vo, **A. Aji**, F. Wang: SATO: A Spatial Data Partitioning Framework for Scalable Query Processing (International Conference on Advances in Geographic Information Systems, SIGSPATIAL 2014)

◇ **A. Aji**, G. Teodoro, F. Wang: Haggis: Turbocharge A MapReduce based Spatial Data Warehousing System with GPU Engine (3rd workshop on Analytics for Big Geospatial Data, BigSpatial 2014)

◇ X. Chen, H. Vo, **A. Aji**, F. Wang: High Performance Integrated Spatial Big Data Analytics. (3rd workshop on Analytics for Big Geospatial Data, BigSpatial 2014)

◇ H. Vo, D. Teng, Y. Liang, **A. Aji**, J. Kong, F. Wang: A MapReduce Based High Performance Whole Slide Image Analysis Framework in the Cloud. (Pathology Informatics 2014)

◇ **A. Aji**, X. Sun, H. Vo, J. Saltz, F. Wang: Demonstration of Hadoop-GIS: A Spatial Data Warehousing System Over MapReduce. (International Conference on Advances in Geographic Information Systems, SIGSPATIAL 2013)

◇ **A. Aji**, F. Wang, H. Vo, R. Lee, Q. Liu, X. Zhang, J. Saltz: Hadoop-GIS: A High Performance Spatial Data Warehousing System Over MapReduce. (International Conference on Very Large Databases, VLDB 2013)

◇ X. Qi, D. Wang, J. Montes, I. Rodero, T Pan, **A. Aji**, L. Cooper, F. Xing, M. Parashar, D. Foran, L. Yang: Exploring Online Nuclear Segmentation on Large Fluorescence Brain Tumor Images using CometCloud. (International Workshop on High Performance Computing for Biomedical Image Analysis, HPC-MICCAI 2013)

◇ **A. Aji**, F. Wang, J. Saltz: Towards Building A High Performance Spatial Query System for Large Scale Medical Imaging Data. (International Conference on Advances in Geographic Information Systems, SIGSPATIAL 2012)

◇ **A. Aji**, Q. Liu, F. Wang, T. Kurc, J. Saltz: MIGIS – A High Performance Query System for Pathology Imaging Analytics. (Pathology Informatics 2012)

◇ **A. Aji**: High Performance Spatial Query Processing for Large Scale Scientific Data. (SIGMOD/PODS PhD Symposium, 2012)

◇ **A. Aji**, Y. Wang, E. Agichtein, E. Gabrilovich: Using the Past to Score the Present: Extending Term Weighting Models with Revision History Analysis. (International Conference on Information and Knowledge Management, CIKM 2010)

◇ **A. Aji**, E. Agichtein: The "Nays" Have It: Exploring Effects of Sentiment in Collaborative Knowledge Sharing. (Computational Linguistics in a World of Social Media, NAACL 2010)

◇ **A. Aji**, E. Agichtein: Deconstructing Interaction Dynamics in Knowledge Sharing Communities. (International Conference on Social Computing, Behavioral Modeling and Prediction, SBP 2010)

◇ M. Theobald, **A. Aji**, R. Schenkel: TopX 2.0 at the INEX 2009 Ad-Hoc and Efficiency Tracks. (INitiative for the Evaluation of XML Retrieval, INEX 2009)

Other Publications

◇ F. Wang, **A. Aji**, H. Vo: High Performance Spatial Queries for Spatial Big Data: from Medical Imaging to GIS. (SIGSPATIAL Newsletter 2015)

◇ F. Wang, **A. Aji**, Q. Liu, J. Saltz: Hadoop-GIS: A High Performance Query System for Analytical Medical Imaging with MapReduce. (Emory University, 2011)

Patents

○ Calculating and Using Out-of-Collection Probabilities For Improved Entity Recognition (US Patent Filed by Microsoft in 2012)

Awards

Best Presentation Award, Pathology Informatics 2012

Travel Awards and Fellowships: SBP 2010, SIGMOD 2012, Pathology Informatics 2012, SIGSPATIAL 2012, VLDB 2013, SIGSPATIAL 2013

Fellowship Award from Laney Graduate School of Arts & Science, Emory University 2008

Distinguished Student Award , University of Science & Technology Beijing 2006

Excellence in Academic Achievement, University of Science & Technology Beijing 2004,2005

Scholarship of Second Honor, University of Science & Technology Beijing 2002-2005

Grants

Assisted in application of a NSF Career Award with the total amount of $445,000.

Teaching Experience

Instructor for 3 computer science undergraduate courses.

TA for 2 undergraduate courses and 2 graduate courses.

Professional Services

PC Member SIGSPATIAL 2014/2015, SIGSPATIAL PhD Symposium 2014, GEOProcessing 2015

Invited Reviewer JCST, CMPB, IJGIS, CCPE

External Reviewer SIGSPATIAL 2013, BigSpatial 2013, VLDB 2013/2014/2015, IEEE BigData 2014.

References