

Jia Jianfeng

jianfeng.jia@gmail.com (949) 678-9893

JOB OBJECTIVE

Seeking a *Summer Internship* position.

WORK EXPERIENCE

Software Engineer Intern SRCH2.com, Irvine, US Jun.2013 - Sept.2013

Implementing a Java SDK for the C++ search engine library. Implementing a Android search app by this SDK.

Research Associate Sogou.com, Beijing, China Jul.2008 - Jul.2012

Working at *Chinese Input Method Editor* (IME) Research Group, working on the Terabytes scale of data to improve the precision of IME product.

EDUCATION

Ph.D student, Computer Science University of California, Irvine, CA, US Sept.2012 - Present

Research interest: *Parallel Computing, Large scale data processing*

M.S. Computer Science Xiamen University, China Sept.2005 - Jul.2008

B.S. Computer Science Xiamen University, China Seps.2001 - Jul.2005

SKILLS

C++, Java, JNI, Android, Python, Hadoop, Hbase, Pig.

PROJECTS

Genome assembling project using Hyracks UCI

Using the Hyracks platform(Parallel data processing system competing with Hadoop) to build the genome graph which contains billions of nodes. And we achieved 30% performance improvement.

Feedback Data flow System of IME using HBase and Pig Sogou.com

Building the automatic feedback data processing system using HBase to store 30G data per day, 15T data totally. Using Pig to analyze and explore the global user behavior and also keeping track of single user's daily statistical features.

Large-scale Language Model(LM) for Cloud IME Sogou.com

Building the automatic process of LM updating weekly from the 500G new corpus using Hadoop platform. Building the decoder for trigram LM and the re-rank model to improve the precision, which is 3% higher than competitors' products.

Automatic New Word Detection Sogou.com

Devising a novel approach of New Word Detection system based on Entropy-loss theory on Hadoop platform.

Dependency Treelet Based Chinese-to-English SMT System Xiamen University

Devising two dependency structure based statistical machine translation system prototypes.