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, Ruby, Python, Hadoop,Hbase,Pig.

PROJECTS

Android SDK for SRCH2 C++ library   SRCH2.com

Developing a Java SDK for the C++ search engine library using JNI and Android NDK. We implemented an error-tolerant search app using this SDK.

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.

PATENTS

A Input Method In The Hardware Device, **Jianfeng Jia**, Yanfeng Wang, Yang Zhang, CN102087550A public on Jun.8, 2011

A Method And A System Providing New Words And Hot Words, **Jianfeng Jia**, Yang Zhang, Yanfeng Wang, CN102163198A, public on Aug.24, 2011

HONORS AND AWARDS

Rhinoceroses prize for building the trigram and re-rank model for cloud IME, Sogou Research, 2010

Rhinoceroses prize for innovation on improve 20% precision rate on the long sentence test set of desktop IME, Sogou Research,2009

First-Class Scholarship, Xiamen University, 2005

First-Class Scholarship, Xiamen University, 2004

PUBLICATIONS

Jianfeng Jia, *The Application of Statistical Language Model in Sogou Pinyin Input Method Editor*, Journal of Chinese Association for Artificial Intelligence 2011.vol.1 (4).

Jianfeng Jia, Xiaodong Shi, Xingbang Lai, *HMM-based Chinese Pinyin Input Method*, J. Modern Computer 2008. (4) 4-6.

Xiaodong Shi, Yidong Chen, **Jianfeng Jia**, *Dependency-Based Chinese-English Statistical Machine Translation*, Conference on Intelligent Text Processing and Computational Linguistics (CICLing) 2007, Mexico City, Mexico.

Jianfeng Jia, Xiaodong Shi, Yu Chen, *Shift-Reduce Algorithm and Structure Model Based Dependency Statistical Parser*, International Chinese Computing Conference (ICCC) 2007, Wuhan, China.