Guanyu Wang

(412)230-7108; guanyuw@andrew.cmu.edu

Education

Carnegie Mellon University Current GPA: 3.92 Sept 2012 - Dec 2013
MS in SCS, specialization in Very Large Information Systems, School of Computer Science

Tsinghua University GPA: 92% Sept 2010 - July 2012
M.S. in Computer Science

University of Science and Technology Beijing GPA: 91%, Rank: 2nd/173 Sept 2006 - July 2010
B.E. in Computer Science

Projects

Slot Filling with Scalable Machine Learning

– (Java, C++) - RA, advised by **Prof. William W. Cohen**, CMU May 2013 - Dec 2013,

- Proposed an algorithm called Focus-Entity Slot Filling algorithm for Slot Filling task, which inputs a single "focus entity" e_f and a set of relations $\mathcal{R}$, then finds all other entities that the focus entity is related to via some relation in $\mathcal{R}$ from a large given unstructured natural language corpus.
- Built a efficient system (sub-system of **NELL**, Never-Ending Language Learner) which can retrieve a collection of entities from the target large text corpus to fill in values for predefined slots (relation attributes) for a given entity in a reference Knowledge Base (Knowledge Base Population).

"Who To Follow" Recommendation System for Social Networks (Java) - CMU Feb 2013 - May 2013

- Explored the "who to follow" recommendation problem in Twitter like social networks.
- Incorporated the social network structure with user and item features to provide recommendation by using the personalized random walk.
- Implemented distributed feature extraction and recommendation system with MapReduce framework and GraphChi (process very large graph computations on a single machine).
- Tested system on KDDCUP12 Tencent Dataset, got rank 50th among about 700 teams.

Node Label Prediction with Belief Propagation on Hadoop (Java) - CMU Sep 2012 - Dec 2012

- Implemented and compared BP and FastBP Algorithms on Hadoop for item category prediction on Amazon co-purchasing dataset (and some others like DBLP).
- Researched into various multi-class classification methods based on two-class classifiers.
- Extended the FastBP algorithm from two-class to multi-class by solving an optimization problem derived from original network to approximate the BP solution.

Biology Question-Answering System (Java) - CMU Sep 2012 - Dec 2012

- Learned question-answering system development under the UIMA framework and researched into various QA system techniques with de-noising, synonym-expansion, HMM, etc.
- Improved passage extractor with a sentence-based self-learning strategy from 1.6% (MAP) baseline to 2.5%.
- Implemented document parser with Apache Solr which enhancing extensibility (allowing adjustable boosts and query parsers, etc.) over original one.

Distributed Channel Discovery Algorithm (Matlab and C/C++) - Independent Oct 2011 - Feb 2012

- Modeled the Whitespace Discovery problem with evil blocking as three players' game; and proposed a constant approximation algorithm for the optimal solution.
- Discovered an optimal strategy for the "multi-sender and one receiver" distributed transmission problem using Probabilistic Power Control.

Data aggregation in Wireless Sensor Network (C/C++) - Tsinghua Oct 2010 - June 2011

- Proposed an algorithm with the binomial tree for Connectivity Problem in wireless network which improves the best theoretical result by a $O(\log n)$ factor.
- Implemented the MAC protocol based on proposed algorithm on Mote-Iris (TinyOS).

Publications

1. **Guanyu Wang**, William W. Cohen, "*From Relation Extraction Slot-Filling: Jointly Reasoning About Different Fillers of a Focus Entity*," submitted to CoNLL'14
2. **Guanyu Wang**, Qiang-Sheng Hua and Yuexuan Wang, "*Data Aggregation in Wireless Sensor Networks under the SINR model*," in ADHOC-NOW'12, July, 2012, Belgrade, Serbia.
3. Zhaoquan Gu, **Guanyu Wang** and Yuexuan Wang, "*Path-Loss Fluctuations: Towards Robust Scheduling Algorithms in the SINR Model*," in IEEE MASS'12, October, 2012, Las Vegas, USA.
4. Zhaoquan Gu, **Guanyu Wang**, Qiang-Sheng Hua and Yuexuan Wang, "*Improved Minimum Latency Aggregation Scheduling in Wireless Sensor Networks*," in International Journal of Sensor Networks

Interns and Services

- TA of *Machine Learning* in Carnegie Mellon University — Sep 2013 - Dec 2013
- TA of *Advanced Algorithm* for Pilot CS Class in Tsinghua University — Sep 2011 - Jan 2012
- Research Intern in Lab of Multimedia and HCI, Peking University. — Aug 2010 - Sep 2010

Honors

- 2010-2011 Contribution Award for ITCS, Tsinghua University
- 2010 Outstanding Graduate of Beijing(1%)
- 2006-2008 National Scholarship (Top grade scholarship)