

# Jingwei Zhang

+1(347)821-1086 ◊ iceboal@gmail.com ◊ http://iceboal.com

## EDUCATION

---

- 2013–2014 **MSc in Computer Science, 4.0** (Machine Learning Track)  
*Columbia University, New York*
- 2009–2013 **BSc in Electronic Engineering**  
*Tsinghua University, Beijing*

## WORK EXPERIENCE

---

### 2015 **SDE @Amazon**

- Joined the Feed and Report Platform Gateway team in Amazon Marketplace.

### 2014 **SDE Intern @Amazon**

- *Feed and Report Platform Daily Flash*  
Used various AWS services and visualization techniques to automatically generate daily report of the usage of Amazon Marketplace Feed and Report Platform services, and provide a web portal for users to view and customize their daily report.

### 2012 **Research Intern @THU & Tencent**

- *Unveiling the Unique Features of Tencent Weibo*  
Used Hadoop to process collected data of more than 320M users and 3.7B microblogs on a cluster of 36 servers. Studied Tencent Weibo from both macro and micro levels, discovered interesting difference between Tencent Weibo and Twitter.

## ACADEMIC EXPERIENCE

---

### 2014 **Research project @CS, Columbia U**

- *Fast, Flexible Models for Discovering Topic Correlation across Weakly-Related Collections*  
Developed two probabilistic topic models, Correlated LDA and Correlated HDP, for analyzing topic correlations between large, asymmetric, and potentially weakly-related collections. Used C-LDA to compare over 300k documents in collections of sciences and humanities research from JSTOR.
- *Word semantic representations using Bayesian Probabilistic Tensor Factorization*  
Developed a Bayesian probabilistic tensor factorization model for generating word vector representations and perspective linear transformations from any number of word similarity perspectives. Evaluated the word embeddings with GRE antonym questions, achieved the state-of-the-art performance. Paper presented in EMNLP'14, project will be integrated in IBM Watson.

### 2014- **Research assistant @CS, Columbia U**

- *Role Discovery based on Email Corpus*  
Developed an application of semi-supervised learning that utilize Author Topic Model and graph Laplacian to automate role discovery with a small amount of training data.
- *Mention Network in the Enron Email Corpus*  
We introduced a novel mention network from emails, then demonstrated such network can help predicting the organizational dominance in the Enron corporation. Implemented graph walk and re-ranking algorithm for name disambiguation.
- *Extracting the Social Network of Taliban*  
Used natural language processing technique to perform social event extraction over a large tranche of newspapers and magazines produced by Taliban, aim to construct a mapping of the social network of the Taliban leadership.

## OTHER REPRESENTATIVE PROJECTS

---

- *Acoustic bird species recognition app on iOS device.*
- *Used Machine Learning to help Kinsa Inc. building their theormometer app.*
- *Greedy forwarding algorithm for mobile social networks embedded in hyperbolic spaces.*
- *Music transcription using Gaussian Mixture Model and beat detection.*

## CONTEST AND AWARDS

---

- 2014 **Amazon Global Intern Hackathon (Computer Vision)**  
*Rank 4th / 73 teams*
- 2013 **Kaggle Contest: Multi-label Bird Species Classification - NIPS 2013**  
*Rank 18th*  
Kaggle.com rank: 964th / 136,948 users
- 2010 **Baidu Cup Tsinghua Duishi AI Contest**  
*Rank 2nd*
- 2009 **National Physics Contest for high school students in China**  
*Rank 4th / ~ 500,000 students*

## SCHOLARSHIPS

---

- 2010 **Samsung Scholarship** (cny 3,000)  
*Tsinghua University*
- 2009 **School of Information Sci. and Tech. Scholarship** (cny 20,000)  
*Tsinghua University*

## PUBLICATIONS

---

**Jingwei Zhang**, Aaron Gerow, Jaan Altosaar, James Evans, Richard Jean So. *Fast, Flexible Models for Discovering Topic Correlation across Weakly-Related Collections*. Submitted to ACL 2015.

**Jingwei Zhang**, Jeremy Salwen, Michael Glass, Alfio Gliozzo. *Word Semantic Representations using Bayesian Probabilistic Tensor Factorization*. EMNLP 2014.

Apoorv Agarwal, Adinoyi Omuya, **Jingwei Zhang**, Owen Rambow. *Enron Corporation: Youre the Boss if People Get Mentioned to You*. SocialCom 2014.

**Jingwei Zhang**. *Greedy forwarding for mobile social networks embedded in hyperbolic spaces*. SIGCOMM 2013: 555–556, poster.

Daifeng Li, **Jingwei Zhang**, Golden Guo-zheng Sun, Jie Tang, Ying Ding, Zhipeng Luo. *What is the Nature of Chinese MicroBlogging: Unveiling the Unique Features of Tencent Weibo*. arXiv:1211.2197.

## SKILLS

---

- GOOD LEVEL    Java, Python, R, MATLAB, C/C++, Hadoop, SQL, L<sup>A</sup>T<sub>E</sub>X, Vim
- BASIC LEVEL    Linux, Shell script, Git, Cython, AWS, d3.js, angular.js