

# Jie Cao

## RESEARCH EXPERIENCE

NLP LAB, University of Utah, SLC 2015 - PRESENT  
I work with Prof. Vivek Srikumar. My current research are mainly in the following region:

- Dialogue System  
Our gated self-attentive memory network for knowledge-grounded goal-oriented dialogue ranked top 2 in DSTC7 challenge track 1 (Accepted by AAAI 2019 [5]). Additionally, my research also covers other dialogue applications such as mental health counseling and creative language generation.
- Semantic Parsing, Structured Prediction  
I also worked on structured prediction methods for Abstract Meaning Representation, especially latent alignment and parsing.

Researcher at SOHU RDC LAB, Beijing 2014 - 2015

- Datacenter-scale Hadoop, Spark DBAS, Distributed Machine Learning  
Large Scale data migration (20 PB) with data security constraints, we utilize hadoop distcp and token server to implement cross-region data migration.

CGCL LAB, HUST, Wuhan 2008 - 2012

- Large Scale Statistical Computing, Programming Language  
By porting R language running on JVM to gain the power of JVM-based open-source data infrastructures such as Hadoop, Spark [3].
- Virtualization, Embedded Virtualization  
Cache performance evaluation [1], task-aware mechanism for non-MMU XEN-ARM based on embedded virtualization[2].

## PRIZES & AWARDS

- 2010 VMware Cloud Computing Innovation Cup, Top 50
- 2009 Google Android Innovative Idea Sharing Award
- 2007 "Computer World" Magazine Scholarship
- 2007 Microsoft Imagine Cup
  - Algorithm Challenge, Top 50 World-Wide
  - Visual Gaming Contest, Top 2 in China, 18th World Final
- 2006 HUST ACM Programming Contest, Top 3

## EDUCATION

- 2015 - Present **Doctor of Philosophy**  
COMPUTER SCIENCE  
*University of Utah*
- 2009 - 2012 **Master of Science**  
COMPUTER SCIENCE  
*Huazhong University of Science and Technology (HUST)*
- 2005 - 2009 **Bachelor of Engineering**  
INFORMATION SECURITY  
*Huazhong University of Science and Technology (HUST)*

## TEACHING & ADVISING

- Spring 2018 TA Structured Prediction, U of Utah
- Fall 2016, 2018 TA Machine Learning, U of Utah
- Fall 2008 Lecturer Algorithm and Data Structure, HUST
- 2006-2008 Leader Unique Studio Algorithm Team, HUST

## INDUSTRY EXPERIENCE

NLP Research Intern at WECHAT AI, Palo Alto 2018

- End-to-End Goal-oriented Dialogue System  
Our dialogue system on DSTC7 ranked top2 in DSTC7. Our paper on Gated self attentive memory network get accepted in AAAI 2019.

Dev Leader at ZUN CLUB, Beijing 2014 - 2015

- Micro-Service Architecture, Recommendation system  
My main work is on heterogeneous data integration and hotel recommendation system

Research Developer at BAIDU, Beijing 2012 - 2014

- Baidu Voice Assistant  
Query analysis, Dialogue, Command dispatcher, Arena framework for Baidu Hackathon on voice assistant
- Mobile Search Engine Architecture  
Index building, Anti-spam, Universal Search Architecture with high performance distributed messaging system.
- Mobile Search Anti-Attack Ecosystem  
Speed optimization, ISP Gateway IP detection, many-core flow clean system, anti-attack policy lighttpd plugins

SDE(Intern) at Alibaba, Hangzhou 2010 - 2011

- MySQL K-V Store  
Speedup MySQL in K-V setting by HandleSocket to bypass complex SQL execution plan and original Mysql protocol.
- Distributed Message Pub-Sub System  
Enhanced MySQL Replication based on MySQL binlog parser, extended slave protocol, and Kafka-like distributed messaging system for efficiently collecting.

## PUBLICATIONS

- [1] Jie Cao. "A Cache Performance Evaluation on Xen Virtualization Platform". In: *Undergraduate Thesis*. 2009.
- [2] Jie Cao, Xia Xie, and Jin Hai. "A real-time schedule system on embedded virtualization". Patent ZL201110410689.1 (CN). 2011.
- [3] Jie Cao et al. "JRBridge: A framework of large-scale statistical computing for R". In: *2012 IEEE Asia-Pacific on Services Computing Conference (APSCC)*. IEEE. 2012, pp. 27-34.
- [4] Xijiang Ke et al. "A distributed SVM method based on the iterative MapReduce". In: *2015 IEEE International Conference on Semantic Computing (ICSC)*. IEEE. 2015, pp. 116-119.
- [5] Jie Cao\* Can-Xiang Yan Zuohui Fu Cheng Niu Shuo Sun\* Yik-Cheung Tam\* and Jie Zhou. "End-to-end Gated Self-attentive Memory Network for Dialog Response Selection". In: *AAAI DSTC7 Workshop*. 2019.

🏠 University Village 967, Salt Lake City, UT, 84108  
☎ +1 (801) 448-5203  
✉ jcao@cs.utah.edu  
🌐 <http://www.mlcv.com>