Shadi A. Noghabi

Department of Computer Science
University of Illinois at Urbana–Champaign
3111 Siebel Center for Computer Science
201 N. Goodwin Ave. Urbana, IL 61801
+1 (734) 546-6697
abdolla2@illinois.edu

Research Interests

Cloud Computing, Big Data, Distributed Systems and Edge Computing

Education

2013–present (exp. 5/2018) **Ph.D. in Computer Science** — **University of Illinois at Urbana–Champaign**
- ⋆ Cumulative GPA: 4.0/4.0

Advisor: Prof. Indy Gupta and Prof. Roy Campbell

2009–2013 **B.Sc. in Computer Engineering** — **Sharif University of Technology**

Publications

2017 **Shadi A. Noghabi**, Kartik Paramasivam, Yi Pan, Navina Ramesh, Jon Bringhurst, Indranil Gupta, Roy H. Campbell, ***Samza: Stateful Stream Processing at Scale***, **VLDB'17**

2016 **Shadi A. Noghabi**, Sriram Subramanian, Priyesh Narayanan Sivabalan Narayanan, Gopalakrishna Holla, Mammad Zadeh, Tianwei Li Indranil Gupta, Roy H. Campbell, ***Ambry: LinkedIn's Scalable Geo-Distributed Object Store***, **SIGMOD'16**

2016 Tianlong Yu, **Shadi A. Noghabi**, Shachar Raindel, Hongqiang Harry Liu, Jitu Padhye, Vyas Sekar, **FreeFlow: High Performance Container Networking**, **HotNets'16**

2016 **Shadi A. Noghabi**, Roy Campbell, Indranil Gupta, **Building a Scalable Distributed Online Media Processing Environment**, PhD workshop **VLDB'16**

2016 Sayed Hadi Hashemi, **Shadi A. Noghabi**, John Bellessa, Roy Campbell, ***Toward Fabric: A Middleware Implementing High-level Description Languages on a Fabric-like Network***, ANCS'16

2013 Mayank Pundir, John Bellessa, **Shadi A. Noghabi**, Cristina L. Abad, Roy H. Campbell, ***Towards Enabling Cooperation Between Scheduler and Storage Layer to Improve Job Performance***, Parallel Data Storage Workshop (PDSW'13 Poster Session)

Technical Reports

2016 Sayed Hadi Hashemi, **Shadi A. Noghabi**, William Gropp, ***Performance Modeling of Distributed Deep Neural Networks***, arXiv:1612.00521

2015 **Shadi A. Noghabi**, Read Sprabery, John Bellessa, Mohammad Ahmad, Indranil Gupta, Roy H. Campbell, ***Real Time Adaptive profiling in Storm Topologies***, UIUC, Technical Report.

2014 Mayank Pundir, Cristina L. Abad, **Shadi A. Noghabi**, Indranil Gupta, John Bellessa, Roy H. Campbell, ***Using Context to Improve Performance of Cloud Stacks*** , UIUC, Technical Report.

2012 **Shadi A. Noghabi**, Sahel Sharifi-Moghadam, Reza Entezari-Maleki, Ali Movaghar, ***New Model for Grid Task Scheduling Based on Priorities and Deadlines***, Performance and Dependability Lab, Sharif University of Technology, Technical Report.

2012 **Shadi A. Noghabi**, Sahel Sharifi-Moghadam, Reza Entezari-Maleki, Ali Movaghar, ***A Communication Cost Aware Scheduling Algorithm for Heterogeneous Environments***, Performance and Dependability Lab, Sharif University of Technology, Technical Report.

Honors and Awards

2017 **Microsoft Research Dissertation Grant** – a $20,000 grant awarded based on the technical merit and impact of the proposed dissertation research.

2017 **Mavis Future Faculty Fellowship** – given by the College of Engineering to students showing promise mid/late career, helping to prepare toward an academic position.

2017 **Tapia Scholarship**, Tapia Confernce'17.

2016 Student Travel Grants for: **SIGMOD'16**, **USENIX ATC'16**, and **ANCS'16**

2016 Grad Cohort Workshop – **CRA Women Scholorship**.

2016 **CS @ ILLINOIS Grace Hopper Grant**, UIUC (*unable to attend*).

2014-2016 Selected to join the **Honor Society of Phi Kappa Phi** for 3 consecutive years.

2015 Selected as **"Active Member"** in **Women in Computer Science (WCS)** association, UIUC

2009–2013 **Ranked in top 5%** based on Cumulative GPA among about 120 students of the department. Class of 2013 students. Recipient of **Honorary Admission for Graduate Study**, Department of Computer Engineering, Sharif University of Technology

2012 **Ranked 7^{th}** in Nationwide Graduate Entrance Qualification Exam (*Konkoor* for graduate study) among more than 17,000 participants, Iran

Research Experiences

Jan'17–present **Microsoft Research**, Mobility and Networking Research Team *Research Scientist*

Developing an Edge computing framework for IoT applications, where developers write IoT applications in a simple API, and the code transparently and optimally gets deployed across the device, Edge and the Cloud.

Jun'16–Sep'16 **Microsoft Research**, Mobility and Networking Research Team *Research Intern*

With the emerge of Cloudlets, with heterogenous while limited hardware specifications, along with wide diversity amongst jobs (resource, bandwidth and latency sensitive), job scheduling becomes very chalenging. In this project, I developed a scalable end-to-end scheduling mechanism hiding resource hetrogenity while optimally scheuling diverse jobs from many users.

Sep'15–present **LinkedIn Corp.** Data Infrastructure Team *Research Software Engineer*

- **Selecting Storage for Stream Processing:** Studying and categorizing storage options for real-world Stream applications and selecting the best storage for each category of applications.
- **Stream Processing Benchmark:** Developing an general puporose stream processing benchmark evaluating various system performance aspects.
- **State in Apache Samza:** Developed fault-tolerant state hanaling at large scale (100s of TBs for a single job) in an unified Lambda-less fashion.
- **Ambry:** Benchmarked and Evalueated Ambry, LinkedIn's geo-distributed object store, serving all media objects across more than 400 million users for over 2 years.

May'15–Aug'15 **LinkedIn Corp.** Data Infrastructure Team, Samza Project *Software Engineering Intern*

I worked on Auto-scaling Apache Samza, LinkedIn's stream processing engine. Samza runs a job on a number of containers, however, currently the user has to specify this number by try-and-error approach. The goal of my project was to remove this burden from the user by having a system that automatically scales out/in.

May'14–Aug'14 **LinkedIn Corp.** Data Infrastructure Team, Ambry Project *Software Engineering Intern*

Evergrowing number of media objects that rarely get deleted necessate continous cluster expansions which in turn create load imbalance. In this project I worked on rebalancing Ambry, LinkedIn's geo-dostibuted object store, with minimum data movement. My approach improved IOPS by 6-10× and storage imbalance by 9-10×.

Aug'13–present **University of Illinois at Urbana-Champaign** *Research Assistant*

- **Real Time Adaptive Profiling in Storm Topologies:** Developed a dynamic profiling engine that runs within Storm and generates improved topologies, optimizing for throughput and latency.
- **High-level Description Languages on a Fabric Network:** Developed a middleware layer for implementing policies and behaviors from high-level network descriptions on top of a Fabric-like network.
- **Using Context to Improve Performance of Cloud Stacks:** Addressing the gap in general-purpose cluster management substrates, such as YARN and Mesos, due to lack of support for passing contextual information in APIs.
- **Mimesis Namespace Generetor:** Developed a namespace generator creating large and realistic hierarchical namespaces, while preserving the input distributions. Source code available on github.
- **Scheduled Caching: Memory Locality with the Help of Scheduler:** Developed a Scheduled Caching technique that leverages information available to the job scheduler by sending pre-fetching hints to the storage layer.

Teaching Experiences

Spring'17 **CS 498: Cloud Computing Applications**, Head TA of the course.

Fall'15 **Cloud Computing Applications**, Coursera Course with more than **9,000** students .

Spring'12&'13 **Theory of Machines and Languages**.

Fall'11 – Fall '12 **Data Structures and Algorithms**.

Fall'12 **Computer Structure and Language**.

Spring'12 **Computer Architecture**.

Spring'11 **Advanced Programming**.

Technical Skills

Development Java, C/C++, C#, Scala, Python, CUDA, Bash

Libraries MATLAB, gnuplot, MPI

Tools/Cloud Storm, Samza, Kafka, MapReduce, Spark, HBase, MySQL, Giraph, familiar with Mahout