Neil Shah

Research Scientist
Snap Inc.
Santa Monica, CA

Email: nshah[at]snap[dot]com
Last Update: July 5, 2019

Education

- **Ph.D**: Computer Science, Carnegie Mellon University, August 2013 - October 2017
  Advisor: Prof. Christos Faloutsos.
- **M.S**: Computer Science, Carnegie Mellon University, August 2013 - May 2017
- **B.S**: Computer Science (Minor in Mathematics), North Carolina State University, August 2010 - May 2013
  GPA: **4.0/4.0 - rank 1**, *Summa Cum Laude* with *Honors*

Positions

- **Snap Inc.**, Research Scientist.
  My work spans computational user modeling and user understanding, especially in the contexts of misbehavior, abuse and fraud on web platforms.
  December 2017 - present
- **Carnegie Mellon University**, Graduate Researcher.
  I worked in the Computer Science Department, on algorithms and applications for anomaly detection in large social graphs.
  August 2013 - October 2017
- **Twitch**, Visiting Researcher.
  I worked on anti-abuse technologies as a member of the Science team.
  January 2016 - May 2016
- **Microsoft Research - Redmond**, Research Intern.
  I worked on improving metrics and methods for measuring research impact for Microsoft Academic Search.
  June 2015 - August 2015
- **Lawrence Livermore National Laboratory**, Research Intern.
  I worked on developing algorithms to automatically identify patterns and anomalies in time-evolving graphs.
  June 2014 - August 2014
- **IBM - Silicon Valley**, Software Intern.
  I worked in the IBM BigInsights group, with a focus on indexing and analytics of system log data.
  May 2012 - January 2013
- **North Carolina State University**, Undergraduate Researcher.
  I worked in the Department of Computer Science on compressing and indexing large scientific datasets.
  June 2009 - April 2013

Awards & Distinctions

- ACM SIGCHI Best Research Paper Honorable Mention Award, 2019
- Symantec Graduate Research Fellowship Finalist, 2017
- ACM SIGKDD Best Research Paper Award, 2016
- National Science Foundation Graduate Research Fellowship, 2013
- North Carolina State University College of Engineering Senior Award for Scholarly Achievement, 2013
- North Carolina State University Department of Computer Science Senior Faculty Scholar, 2012
- National Science Foundation Research Experience for Undergraduates Grant, 2011
- North Carolina State University College of Engineering Dean's Research Assistantship, 2011

- North Carolina State University Caldwell Fellowship, 2011
- Coca-Cola Scholarship, 2010
- Zinch Scholarship, 2010
- National Merit Scholarship, 2010
- CompTIA Information Technology Merit Award, 2010
- 2nd place, National Siemens Competition in Math, Science and Technology, 2009
- 1st place, Regional Siemens Competition in Math, Science and Technology, 2009

Publications

Refereed Conference Publications

27. Hemank Lamba, **Neil Shah**. *Modeling Dwell Time Engagement on Visual Multimedia*, KDD 2019.
26. Hana Habib, **Neil Shah**, Rajan Vaish. *Impact of Contextual Factors on Public Snapchat Sharing*, CHI 2019.
25. Shreya Jain, Dipankar Niranjan, Hemank Lamba, **Neil Shah**, Ponnurangam Kumaraguru. *Characterizing and Detecting Livestreaming Chatbots*, ASONAM 2019
24. Srijan Kumar, **Neil Shah**. *False Information on the Web and Social Media: A Survey*, Social Media Analytics: Advances and Applications, CRC Press 2018.
23. Gisel Batista Guacho, Sara Abdali, **Neil Shah**, Evangelos Papalexakis. *Semi-Supervised Content-based Detection of Misinformation via Tensor Embeddings*, ASONAM 2018.
22. Nikhil Gupta, Dhivya Eswaran, **Neil Shah**, Leman Akoglu, Christos Faloutsos. *Beyond Outlier Detection: LookOut for Pictorial Explanation*, ECML-PKDD 2018.
21. Rohan Kumar, Mohit Kumar, **Neil Shah**, Christos Faloutsos. *Did We Get It Right? Predicting Query Performance in E-commerce Search*, SIGIR eCom 2018.
20. **Neil Shah**, Hemank Lamba, Alex Beutel and Christos Faloutsos. *The Many Faces of Link Fraud*, ICDM 2017.
19. Da-Cheng Juan, **Neil Shah**, Mingyu Tang, Zhiliang Qian, Diana Marculescu, Christos Faloutsos. *M3A: Model, MetaModel, and Anomaly Detection in Web Searches*, DSAA 2017.
18. **Neil Shah**. *FLOCK: Combating Astroturfing on Livestreaming Platforms*, WWW 2017.
17. **Neil Shah**, Alex Beutel, Bryan Hooi, Leman Akoglu, Stephan Gunnemann, Disha Makhija, Mohit Kumar, Christos Faloutsos. *EdgeCentric: Anomaly Detection in Edge-Attributed Networks*, ICDM DMCS 2016.
16. Bryan Hooi, Hyun Ah Song, Alex Beutel, **Neil Shah**, Kijung Shin, Christos Faloutsos. *FRAUDAR: Bounding Graph Fraud in the Face of Camouflage*, KDD 2016.
15. Yike Liu, Tara Safavi, **Neil Shah**, Danai Koutra. *Reducing Million-Node Graphs to a Few Structural Patterns: A Unified Approach*, KDD MLG 2016.
14. Yike Liu, **Neil Shah**, Danai Koutra. *An Empirical Comparison of the Sumarization Power of Graph Clustering Methods*, NIPS NSIS 2015.
13. **Neil Shah**, Yang Song. *s-Index: Towards Building Better Metrics for Quantifying Research Impact*, arXiv preprint (Submitted).
12. Bryan Hooi, **Neil Shah**, Alex Beutel, Stephan Gunnemann, Leman Akoglu, Mohit Kumar, Disha Makhija, Christos Faloutsos. *BIRDNEST: Bayesian Inference for Ratings-Fraud Detection*, SDM 2016.
11. **Neil Shah**, Danai Koutra, Tianmin Zou, Brian Gallagher, Christos Faloutsos. *TimeCrunch: Interpretable Dynamic Graph Summarization*, KDD 2015.
10. Maria Giatsoglou, Despoina Chatzakou, **Neil Shah**, Alex Beutel, Stephan Guenneman, Christos Faloutsos, Athena Vakali. *ND-Sync: Detecting Synchronized Fraud Activities*, PAKDD 2015.
9. Maria Giatsoglou, Despoina Chatzakou, **Neil Shah**, Christos Faloutsos, Athena Vakali. *Retweeting Activity on Twitter: Signs of Fraud*, PAKDD 2015.
8. **Neil Shah**, Alex Beutel, Brian Gallagher, Christos Faloutsos. *Spotting Suspicious Link Behavior with fBox: An Adversarial Perspective*, ICDM 2014.
7. **Neil Shah**, Eric Schendel, Saurabh Pendse, Sriram Lakshminarasimhan, Terry Rogers, Nagiza Samatova. *Improving I/O Throughput with PRIMACY: Preconditioning ID-Mapper for Compressing Incompressibility*, CLUSTER 2012.

6. Eric Schendel, Ye Jin, **Neil Shah**, Jackie Chen, Choong-Seock Chang, Seung-Hoe Ku, Stephane Ethier, Scott Klasky, Robert Latham, Robert Ross, Nagiza Samatova. *ISOBAR Preconditioner for Effective and High-throughput Lossless Data Compression*, ICDE 2012.
5. Isha Arkatkar, John Jenkins, Sriram Lakshminarasimhan, **Neil Shah**, Eric Schendel, Stephane Ethier, Choong-Seock Chang, Jackie Chen, Hemant Kolla, Scott Klasky, Robert Ross, Nagiza Samatova. *Analytics-driven Lossless Data Compression for Rapid In-situ Indexing, Storing and Querying*, DEXA 2012.
4. Ye Jin, Sriram Lakshminarasimhan, **Neil Shah**, Zhenhuan Gong, Choong-Seock Chang, Jackie Chen, Stephane Ethier, Hemant Kolla, Seung-Hoe Ku, Scott Klasky, Robert Latham, Robert Ross, Karen Schuchardt, Nagiza Samatova. *S-preconditioner for Multi-fold Data Reduction with Guaranteed User-controlled Accuracy*, ICDM 2011.
3. Sriram Lakshminarasimhan, **Neil Shah**, Stephane Ethier, Scott Klasky, Robert Latham, Robert Ross, Nagiza Samatova. *Compressing the Incompressible with ISABELA: In-situ Reduction of Spatio-Temporal Data*, EuroPar 2011.
2. **Neil Shah**, Yekaterina Shpanskaya, Choong-Seock Chang, Seung-Hoe Ku, Anatoli Melechko, Nagiza Samatova. *Automatic and Statistically Robust Spatio-temporal Detection and Tracking of Fusion Plasma Fronts*, SciDAC 2010.
1. Paul Breimyer, Guruprasad Kora, William Hendrix, **Neil Shah**, Nagiza Samatova. *pR: Automatic Parallelization of Data-parallel Statistical Computing Codes for R in Hybrid Multi-node and Multi-core Environments*, IADIS 2009.

Refereed Journal Publications

4. Yike Liu, Tara Safavi, **Neil Shah**, Danai Koutra. *Reducing Large Graphs to Small Supergraphs: A Unified Approach*, Social Network Analysis and Mining 2018.
3. **Neil Shah**, Danai Koutra, Lisa Jin, Tianmin Zou, Brian Gallagher, Christos Faloutsos. *On Summarizing Large-Scale Dynamic Graphs*, Data Engineering Bulletin 2017.
2. Danai Koutra, **Neil Shah**, Joshua T. Vogelstein, Brian Gallagher, Christos Faloutsos. *DeltaCon: A Principled Massive-Graph Similarity Function with Attribution*, Transactions on Knowledge and Data Discovery 2015.
1. Sriram Lakshminarasimhan, **Neil Shah**, Stephane Ethier, Scott Klasky, Robert Latham, Robert Ross, Nagiza Samatova. *ISABELA for Effective In-situ Compression of Scientific Data*, Concurrency and Computation: Practice and Experience 2011.

Book Chapters

3. Srijan Kumar, **Neil Shah**. *False Information on the Web and Social Media: A Survey*, Social Media Analytics: Advances and Applications.
2. **Neil Shah**. *Introduction to R*, Practical Graph Mining with R.
1. Kanchana Padmanabhan, Sriram Lakshminarasimhan, Zhenhuan Gong, John Jenkins, **Neil Shah**, Eric Schendel, Isha Arkatkar, Robert Ross, Scott Klasky, Nagiza Samatova. *In-situ Analysis in Support of Exploratory Scientific Discovery in Data-Intensive Science*, Data-Intensive Science.

Academic Experience

Service and Mentoring

- PC Chair for WSDM Cup (2020)
- PC Chair for ASONAM Industrial Track (2019)
- PC Chair for WWW Cybersafety Workshop (2019)
- Invited Keynote Speaker for WWW Cybersafety Workshop (2018)
- Invited Keynote Speaker for KDD Outlier Detection De-constructed Workshop (2018)
- Invited Keynote Speaker for ECML-PKDD PhD Forum (2018)
- Session Chair for WWW "Security and Privacy" track (2018)
- Session Chair for ICDM "Social" track (2016)

- Program Committee for KDD (2019)
- Program Committee for WWW (2015, 2018)
- Program Committee for SDM (2018)
- Program Committee for CIKM (2017)
- Program Committee for ICDM Demo Session (2019)
- Program Committee for WSDM "Misinformation and Misbehavior Mining on the Web" Workshop (2018)
- Program Committee for ICDM Demo Session (2018)
- Program Committee for KDD "Mining and Learning with Graphs" Workshop (2017, 2018, 2019)
- Program Committee for WSDM "Heterogeneous Networks Analysis and Mining" Workshop (2017, 2018)
- Program Committee for ICDM PhD Forum (2017)
- Reviewer for CSCW (2019)
- Reviewer for CHI (2019)
- Reviewer for Springer "Data Mining and Knowledge Discovery" Journal (2018)
- Reviewer for ACM "Transactions on Social Computing" Journal (2018)
- Reviewer for ACM "Transactions on Knowledge Discovery from Data' Journal (2018, 2019, 2020)
- Reviewer for ACM "Transactions on Knowledge and Data Engineering" Journal (2016, 2017)
- Reviewer for WISE (2014)
- Reviewer for IPDPS (2011)
- Mentored Dipankar Niranjan (BS student at IIIT Delhi)
- Mentored Shreya Jain (BS student at IIIT Delhi)
- Mentored Mr. Himel Dev (intern at Snap Research)
- Mentored Mr. Hamed Nilforoshan (intern at Snap Research)
- Mentored Mrs. Hana Habib (intern at Snap Research)
- Mentored Mr. Rohan Kumar (visiting CS student at CMU)
- Mentored Mr. Hemank Lamba (ISR PhD student at CMU)
- Mentored Ms. Qicheng Huang (EE PhD student at CMU)
- Mentored Ms. Chenlei Fang (EE PhD student at CMU)
- Mentored Mr. Tianmin Zou (CS MS student at CMU)

Funding

- Contributed towards Flipkart faculty grant (with CMU: Bryan Hooi, Dhivya Eswaran, Christos Faloutsos)
- Contributed towards Wharton Customer Analytics Initiative proposal "Fraud Detection through Mining Dynamic Behavior for Group Anomalies" (with CMU: Alex Beutel and Christos Faloutsos)
- Contributed towards PNC Center for Financial Services proposal PF15003: "Fraud Detection in Financial Data" (with CMU: Alex Beutel and Christos Faloutsos)
- Contributed towards proposal DOE-NNSA-30788.1.1990222 "Quantifying Network Changes" (with CMU: Danai Koutra and Christos Faloutsos)
- Contributed towards proposal NSF IIS-1028746 "Collaborative Research: Understanding Climate Change: A Data Driven Approach" (with NCSU: Nagiza Samatova and Fredrick Semazzi)

Teaching

- Guest Lecture "A Foray into Graph Mining" for USC CSCI-699 Introduction to Information Extraction, by Prof. Xiang Ren
- Guest Lecture "Graph Mining for Fraud Detection" for CMU 15-300 Research and Innovation in Computer Science, by Prof. Todd Mowry
- Teaching Assistant for CMU 15-300 Research and Innovation in Computer Science, by Prof. Todd Mowry
- Teaching Assistant for CMU 15-826 Multimedia Databases and Data Mining, by Prof. Christos Faloutsos

Technical Skills

- Cloud Platforms: GCP, AWS

- Programming Languages/Tools: Python, Java, C, C++, x86 Assembly
- Web Languages/Tools: SQL, HTML, PHP, JavaScript, CSS, WordPress
- Engineering Tools: Matlab, R
- Typesetting Tools: LaTeX, Microsoft Office, LibreOffice
- Source Code Management Tools: Git, Subversion
- Operating Systems: Microsoft Windows, Ubuntu Linux, Mac OSX

Graduate Coursework

Carnegie Mellon University

- *Advanced and Distributed Operating Systems* with Prof. Dave Andersen,
- *Multimedia Databases and Data Mining* with Prof. Christos Faloutsos
  Graduate Algorithms with Prof. Manuel Blum
- *Machine Learning* with Profs. Eric Xing and Aarti Singh
- *Randomized Algorithms* with Prof. Bernhard Haeupler
- *Programming Language Semantics* with Prof. André Platzer

North Carolina State University

- *Automated Learning & Data Analysis* with Prof. Nagiza Samatova,
- *Software Engineering* with Prof. Tao Xie,
- *Database Management Concepts & Systems* with Prof. Ting Yu

References

Available upon request.