

# Nan Zhang

Kogod School of Business, American University, Washington DC  
Tel: (202) 885-1765, Email: [nzhang@american.edu](mailto:nzhang@american.edu), Web: <https://robustanalytics.org/nan>

## Research Interests

Data analytics; Information privacy and security; Machine learning in social and behavioral science research.

## Appointments

American University, Washington DC 2018 – Present  
*Professor*  
Kogod School of Business

The Pennsylvania State University 2017 – 2018  
*Professor*  
College of Information Sciences and Technology

National Science Foundation 2016 – 2017  
*Program Director*

Division of Information & Intelligent Systems (CISE/IIS)

Program responsibilities:

- Information Integration and Informatics (III)
  - managing research projects in the areas of data management and data analytics
- Secure and Trustworthy Cyberspace (SaTC)
  - serving as the IIS liaison in the SaTC program
  - managing research projects in privacy and data science
  - facilitating interdisciplinary research on privacy and cybersecurity (16-027: Enabling US-Netherlands Collaboration for Privacy Research, 17-019: Enabling New Collaborations Between CISE and SBE Research Communities)
- Critical Techniques, Technologies and Methodologies for Advancing Foundations and Applications of Big Data Sciences and Engineering (BIGDATA)
  - managing research projects in the areas of data management and data analytics

The George Washington University 2008 – 2017  
*Professor, Associate Professor, and Assistant Professor*  
Department of Computer Science

The University of Texas at Arlington 2006 – 2008  
*Assistant Professor*  
Department of Computer Science and Engineering

## Education

Ph.D., Computer Science, Texas A&M University 2006

B.S., Computer Science, Peking University, China 2001

### *Selected Honors and Awards*

- CACM Research Highlight 2020  
nominated by SIGMOD and selected by the *Communications of the ACM* as one of the “outstanding research articles selected from the broad spectrum of computing research”
- ACM SIGMOD Research Highlight Award 2019  
from ACM SIGMOD, paper “Leveraging Similarity Joins for Signal Reconstruction” selected as one of the best research papers published in 2018 that exemplify core database research
- Best Paper Award 2019  
From American Marketing Association (AMA) Summer Academic Conference, Consumer 360 Track
- Best Paper Nomination 2018, 2019  
from Hawaii International Conference on System Sciences
- Distinguished Member of the Technical Program Committee 2017  
from IEEE International Conference on Computer Communications (INFOCOM)
- Best Paper Nomination 2015  
from IEEE International Conference on Intelligence and Security Informatics (ISI), 1 of 5 nominated
- Distinguished Young Lecturer 2014  
from The International Conference on Web-Age Information Management (WAIM 2014).
- Best Student Paper Award 2013  
from ACM International Conference on Information and Knowledge Management (CIKM 2013). One of the two best papers (along with the best paper award) selected from 1,211 submissions.
- Best Paper Award 2013  
from IEEE International Conference on Communications (ICC 2013), Communication and Information Systems Security Symposium
- Best Paper Award 2010  
from IEEE International Conference on Networking, Architecture, and Storage (NAS)
- CAREER Award 2008  
from the National Science Foundation (NSF). According to NSF, this is the most prestigious award in support of junior faculty who exemplify the role of teacher-scholars through outstanding research, excellent education, and the integration of education and research.

### *Recent Research Grants (2016 - Present)*

1. PI, FAI: Using Machine Learning to Address Structural Bias in Personnel Selection, \$999,177, Fairness in Artificial Intelligence (FAI) Program, jointed sponsored by National Science Foundation and Amazon, 2/1/2021 - 1/31/2024, with Heng Xu (Co-PI), Mo Wang (Co-PI).

2. PI, Theore: Theory-Driven Curation and Reusable Robustness Calculus of Social and Behavioral Sciences (SBS) Claims, \$249,683, Defense Advanced Research Projects Agency (DARPA), Defense Sciences Office (DSO) Next Generation Social Science (NGS2) Program, 6/17/2019 - 7/16/2020, with Heng Xu (Co-PI).
3. Co-I, The Center for Innovation in Intensive Longitudinal Studies, \$1,896,924, National Institute of Health, 9/20/2018 - 8/31/2022, with Sy-miin Chow (PI), Nilam Ram et al. (Co-Is).
4. Co-PI, SaTC: CORE: Medium: Situation-Aware Identification and Rectification of Regrettable Privacy Decisions, \$904,133, National Science Foundation, Secure and Trustworthy Cyberspace (SaTC) Program, 8/1/2018 - 7/31/2022, with Heng Xu (PI).
5. Co-PI, RR: Establishing & Boosting Confidence Levels for Empirical Research Using Twitter Data, \$400,000, National Science Foundation, Science of Science and Innovation Policy (SciSIP) Program, 3/15/2018 - 3/14/2021, with Heng Xu (PI).
6. PI, Intergovernmental Personnel Act (IPA) Assignment, National Science Foundation, \$390,506, 1/11/2016 - 1/10/2018.
7. PI, Analytics-based Interface Transformation for Web Databases, \$899,958.00 (GW Portion: \$156,568), Qatar National Research Fund, 4/1/2015 - 3/31/2018, with Ali Jaoua (Qatar University), Gautam Das (The University of Texas at Arlington).
8. Co-PI, US-Korea Workshop on Security and Privacy for Internet of Things, \$30,000, National Science Foundation, Catalyzing New International Collaborations (CNIC) Program, 3/15/2015 - 2/29/2016, with Hyeong-Ah Choi (PI) and Xiuzhen Cheng (co-PI).
9. Co-PI (PI at GW), Efficient Analytics over Hidden Online Social Networks, \$449,890 (GW Portion: \$208,745), Army Research Office, 12/1/2014 - 11/30/2017, with Gautam Das (PI, The University of Texas at Arlington).
10. Co-PI, Multi-Input Multi-Output (MIMO) Aware Cooperative Dynamic Spectrum Access, \$749,374 (GW Portion: \$300,000), National Science Foundation, Enhancing Access to the Radio Spectrum (EARS) Program, 9/1/2014 - 8/31/2017, with Xiuzhen Cheng (PI, GWU) and Hyeong-Ah Choi (co-PI, GWU).
11. PI, Privacy-Preserving Framework for Publishing Electronic Healthcare Records, \$754,911 (GW Portion: \$269,505), National Science Foundation, Smart and Connected Health (SCH) Program, 1/1/2014-12/31/2016, with Heng Huang (The University of Texas at Arlington), Liam O'Neill (University of North Texas Health Science Center).

### *Selected Publications*

### *Selected Journal Publications*

1. H. Xu and N. Zhang, Implications of Data Anonymization on the Statistical Evidence of Disparity, *Management Science*, Forthcoming.
2. N. Zhang and H. Xu, Reconciling the Paradoxical Findings of Choice Overload Through an Analytical Lens, *MIS Quarterly*, Forthcoming.
3. N. Zhang, M. Wang, and H. Xu, Disentangling Effect Size Heterogeneity in Meta-analysis: A Latent Mixture Approach, *Psychological Methods*, Forthcoming.
4. A. Asudeh, J. Augustine, A. Nazi, S. Thirumuruganathan, N. Zhang, G. Das, D. Srivastava, Scalable Signal Reconstruction for a Broad Range of Applications, *Communications of the ACM*, 64(2), pp. 106-115, 2021.
5. H. Xu, N. Zhang, and L. Zhou, Validity Concerns In Research Using Organic Data, *Journal of Management*, 46(7), pp. 1257-1274, 2020.
6. A. Asudeh, J. Augustine, A. Nazi, S. Thirumuruganathan, N. Zhang, G. Das, D. Srivastava, Scalable Algorithms for Signal Reconstruction by Leveraging Similarity Joins, *The VLDB Journal*, 29, pp. 681-707, 2020.
7. A. Asudeh, J. Augustine, A. Nazi, S. Thirumuruganathan, N. Zhang, G. Das, D. Srivastava, Efficient Signal Reconstruction for a Broad Range of Applications, *SIGMOD Record*, 48(1), pp. 42-49, 2019.
8. H. Xu, N. Zhang, Privacy in Health Disparity Research, *Medical Care*, 57(6), S172-S175, 2019.
9. A. Asudeh, A. Nazi, J. Augustine, S. Thirumuruganathan, N. Zhang, G. Das, D. Srivastava, Leveraging Similarity Joins for Signal Reconstruction, *Proceedings of the VLDB Endowment (PVLDB)*, 11(10), pp. 1276-1288, 2018. (Selected for the [PVLDB Reproducibility Highlight Award](#) in 2018 for its zero-effort reproducibility; the [SIGMOD Research Highlight Award](#) in 2019; and the [Communications of the ACM Research Highlight](#) in 2020)
10. J. Lin, W. Yu, N. Zhang, X. Yang, and L. Ge, Data Integrity Attacks against Dynamic Route Guidance in Transportation-based Cyber-Physical Systems: Modeling, Analysis, and Defense, *IEEE Transactions on Vehicular Technology*. 67(9), pp. 8738-8753, 2018.
11. J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, W. Zhao, A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications, *IEEE Internet of Things Journal*, 4(5), pp. 1125-1142, 2017.
12. L. O'Neill, F. Dexter, N. Zhang, The Risks to Patient Privacy from Publishing Data from Clinical Anesthesia Studies, *Anesthesia & Analgesia*, 122(6), pp. 2017-2027, 2016. ([feature article](#))
13. A. Asudeh, N. Zhang, G. Das, Query Reranking As a Service, *Proceedings of the VLDB Endowment (PVLDB)*, 9(11), pp. 888-899, 2016.
14. A. Asudeh, S. Thirumuruganathan, N. Zhang, G. Das, Discovering the Skyline of Web Databases, *Proceedings of the VLDB Endowment (PVLDB)*, 9(7), pp. 600-611, 2016.

15. H. Yan, Z. Gong, N. Zhang, T. Huang, H. Zhong, J. Wei, Crawling Hidden Objects with kNN Queries, *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 28(4), pp.912-924, 2016.
16. Z. Zhou, N. Zhang, Z. Gong, G. Das, Faster Random Walks By Rewiring Online Social Networks On-The-Fly, *ACM Transactions on Database Systems (TODS)*, 40(4), Article No. 26, 2016.
17. S. Wang, R. Bie, F. Zhao, N. Zhang, and X. Cheng, Security in Wearable Communications. *IEEE Network*, 30(5), pp. 61-67, 2016.
18. X. Pan, Z. Ling, A. Pingley, W. Yu, N. Zhang, K. Ren, X. Fu, Password Extraction via Reconstructed Wireless Mouse Trajectory, *IEEE Transactions on Dependable and Secure Computing (TDSC)*, 13(4), pp. 461-473, 2016.
19. Y. Lu, S. Thirumuruganathan, N. Zhang, G. Das, Hidden Database Research and Analytics (HYDRA) System, *IEEE Data Engineering Bulletin*, 38(3), pp. 84-102, 2015.
20. W. Liu, M. F. Rahman, S. Thirumuruganathan, N. Zhang, G. Das, Aggregate Estimations over Location Based Services, *Proceedings of the VLDB Endowment (PVLDB)*, 8(12), pp. 1334-1345, 2015.
21. M. F. Rahman, W. Liu, S. Thirumuruganathan, N. Zhang, G. Das, Privacy Implications of Database Ranking, *Proceedings of the VLDB Endowment (PVLDB)*, 8(10), pp. 1106-1117, 2015.
22. Z. Zhou, N. Zhang, G. Das, Leveraging History for Faster Sampling of Online Social Networks, *Proceedings of the VLDB Endowment (PVLDB)*, 8(10), pp. 1034-1045, 2015.
23. A. Nazi, Z. Zhou, S. Thirumuruganathan, N. Zhang, G. Das, Walk, Not Wait: Faster Sampling Over Online Social Networks, *Proceedings of the VLDB Endowment (PVLDB)*, 8(6), pp. 678-689, 2015.
24. H. Yan, Z. Gong, N. Zhang, T. Huang, H. Zhong, J. Wei, Aggregate Estimation in Hidden Databases with Checkbox Interfaces, *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 27(5), pp. 1192-1204, 2015.
25. H. Huang, R. C. Chiang, S. Rajasekaran, N. Zhang, Swiper: Exploiting Virtual Machine Vulnerability in Third-Party Clouds with Competition for I/O Resources, *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 26(6), pp. 1732-1742, 2015.
26. H. Li, X. Cheng, K. Li, C. Hu, N. Zhang, W. Xue, Robust Collaborative Spectrum Sensing Schemes for Cognitive Radio Networks, *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 25(8), pp. 2190-2200, 2014.
27. W. Liu, S. Thirumuruganathan, N. Zhang, G. Das, Aggregate Estimation Over Dynamic Hidden Web Databases, *Proceedings of the VLDB Endowment (PVLDB)*, 7(12), 1107-1118, 2014.
28. N. Zhang, C. Li, N. Hassan, S. Rajasekaran, G. Das, On Skyline Groups, *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 26(4), pp. 942-956, 2014.

29. Q. Yang, J. Yang, W. Yu, D. An, N. Zhang, W. Zhao, On False Data Injection Attacks against Power System State Estimation: Modeling and Countermeasures, *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 25(3), pp. 717-729, 2014.
30. S. Thirumuruganathan, N. Zhang, G. Das, Rank Discovery From Web Databases, *Proceedings of the VLDB Endowment (PVLDB)*, 6(13), pp. 1582-1693, 2013.
31. W. Cheng, N. Zhang, X. Cheng, M. Song, and D. Chen, Time-Bounded Essential Localization for Wireless Sensor Networks, *IEEE/ACM Transactions on Networking*, 21(2), pp. 400-412, 2013.
32. C. Sheng, N. Zhang, Y. Tao, and X. Jin, Optimal Algorithms for Crawling a Hidden Database in the Web, *Proceedings of the VLDB Endowment (PVLDB)*, 5(11), pp. 1112-1123, 2012.
33. H. H. Huang, N. Zhang, W. Wang, G. Das, and A. Szalay, Just-In-Time Analytics on Large File Systems, *IEEE Transactions on Computers (TC)*, 61(11), pp. 1651-1664, 2012.
34. X. Fu, N. Zhang, A. Pingley, W. Yu, J. Wang, and W. Zhao, The Digital Marauder's Map: a WiFi Forensic Positioning Tool, *IEEE Transactions on Mobile Computing (TMC)*, 11(3), pp. 377-389, 2012.
35. X. Jin, N. Zhang, A. Mone, and G. Das, Randomized Generalization for Aggregate Suppression Over Hidden Web Databases, *Proceedings of the VLDB Endowment (PVLDB)*, 4(11), pp. 1099-1110, 2011.
36. N. Zhang and W. Zhao, Privacy-Preserving OLAP: An Information-Theoretic Approach, *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 23(1), pp.122-138, 2011.
37. W. Yu, N. Zhang, X. Fu, R. Bettati, and W. Zhao, Localization Attacks to Internet Threat Monitors: Modeling and Countermeasures, *IEEE Transactions on Computers (TC)*, 59(12), pp. 1655-1668, 2010.
38. W. Yu, N. Zhang, X. Fu, and W. Zhao, Self-Disciplinary Worms and Countermeasures: Modeling and Analysis, *IEEE Transactions on Parallel and Distributed Systems (TPDS)*, 21(10), pp. 1501-1514, 2010.
39. N. Zhang and W. Zhao, Privacy Protection Against Malicious Adversaries in Distributed Information Sharing Systems, *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 20(8), pp. 1028-1033, 2008.
40. N. Zhang and W. Zhao, Privacy-Preserving Data-Mining Systems, *IEEE Computer*, 40(4), pp. 52-58, 2007.

*Selected Refereed Conference Publications*

41. A. Asudeh, A. Nazi, N. Zhang, G. Das, H. V. Jagadish, RRR: Rank-Regret Representative, *Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD)*, 2019.

42. Y. D. Gunasekaran, M. F. Rahman, S. Hasani, N. Zhang, G. Das, DBLOC: Density Based Clustering over LOCATION Based Services, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), Demo Paper, 2018.
43. K. Peguero, N. Zhang, X. Cheng, An Empirical Study of the Framework Impact on the Security of JavaScript Web Applications, The Web Conference (WWW), 2018.
44. M. F. Rahman, W. Liu, S. Bin Suhaim, S. Thirumuruganathan, N. Zhang and G. Das, HDBSCAN: Density based Clustering over Location Based Services, Proceedings of the IEEE International Conference on Data Engineering (ICDE), 2017.
45. A. Asudeh, A. Nazi, N. Zhang, G. Das, Efficient Computation of Regret-ratio Minimizing Set: A Compact Maxima Representative, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), 2017.
46. M. F. Rahman, S. Bin Suhaim, W. Liu, S. Thirumuruganathan, N. Zhang, G. Das, ANALOC: Efficient ANALytics over LOCATION Based Services, Proceedings of the IEEE International Conference on Data Engineering (ICDE), Demo Paper, 2016.
47. A. Nazi, S. Thirumuruganathan, V. Hristidis, N. Zhang, G. Das, Answering Complex Queries in an Online Community Network, Proceedings of the International AAAI Conference on Web and Social Media (ICWSM), 2015.
48. W. Liu, S. Bin Suhaim, S. Thirumuruganathan, N. Zhang, G. Das, A. Jaoua, HDBTracker: Aggregate Tracking and Monitoring Over Dynamic Web Databases, 40th International Conference on Very Large Data Bases (VLDB), Demo Paper, 2014.
49. S. Thirumuruganathan, N. Zhang, V. Hristidis, G. Das, Aggregate Estimation Over a Microblog Platform, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), 2014.
50. N. Zhang, G. Das, Exploration and Mining of Web Repositories, The ACM International Conference on Web Search and Data Mining (WSDM), Tutorial, 2014.
51. M. Zhang, N. Zhang, G. Das, Mining a Search Engine's Corpus Without a Query Pool, Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM), 2013. (*Best Student Paper Award*)
52. S. Thirumuruganathan, N. Zhang, G. Das, Breaking the Top-k Barrier of Hidden Web Databases, Proceedings of the IEEE International Conference on Data Engineering (ICDE), 2013.
53. Z. Zhou, N. Zhang, Z. Gong, G. Das, Faster Random Walks By Rewiring Online Social Networks On-The-Fly, Proceedings of the IEEE International Conference on Data Engineering (ICDE), 2013.
54. C. Li, N. Zhang, N. Hassan, S. Rajasekaran, G. Das, On Skyline Groups, Proceedings of the ACM International Conference on Information and Knowledge Management (CIKM), 2012.

55. M. Zhang, N. Zhang, and G. Das, Aggregate Suppression for Enterprise Search Engines, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), 2012.
56. N. Zhang, G. Das, Exploration of Deep Web Repositories, 37th International Conference on Very Large Data Bases (VLDB), Tutorial, 2011.
57. X. Jin, N. Zhang, G. Das, Attribute Domain Discovery for Hidden Web Databases, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), 2011.
58. M. Zhang, N. Zhang, G. Das, Mining Enterprise Search Engine's Corpus: Efficient Yet Unbiased Sampling and Aggregate Estimation, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), 2011.
59. X. Jin, A. Mone, N. Zhang, G. Das, MOBIES: Mobile-Interface Enhancement Service for Hidden Web Database, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), Demo Paper, 2011.
60. X. Jin, M. Zhang, N. Zhang, and G. Das, Versatile Publishing for Privacy Preservation, Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), 2010.
61. A. Dasgupta, X. Jin, B. Jewell, N. Zhang, and G. Das, Unbiased estimation of size and other aggregates over hidden web databases, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), 2010.
62. A. Dasgupta, N. Zhang, G. Das, and S. Chaudhuri, Privacy Preservation of Aggregates in Hidden Databases: Why and How?, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), 2009.
63. A. Maiti, A. Dasgupta, N. Zhang, and G. Das, HDSampler: Revealing data behind web form interfaces, Proceedings of the ACM SIGMOD International Conference on Management of Data (SIGMOD), Demo Paper, 2009.
64. A. Dasgupta, N. Zhang, and G. Das, Leveraging COUNT Information in Sampling Hidden Databases, Proceedings of the 25th IEEE International Conference on Data Engineering (ICDE), 2009.
65. N. Zhang and W. Zhao, Distributed Privacy-Preserving Information Sharing, Proceedings of the 31st International Conference on Very Large Data Bases (VLDB), Pages 889-900, 2005.
66. N. Zhang, S. Wang, and W. Zhao, A New Scheme on Privacy-Preserving Data Classification, Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD), Pages 374-383, 2005.

### *Patents*

1. Nan Zhang, Gautam Das, Just-in-time analytics on large file systems and hidden databases, USPTO patent 9,244,976, January 26, 2016.



2. H. Howie Huang, Nan Zhang, Gautam Das, Alex Szalay, Just-In-Time Analytics on Large File Systems, USPTO patent 9,244,975, January 26, 2016.

### *Selected Professional Activities*

1. [Government Service] Co-Chair for Federal Interagency Working Group on Privacy Research and Development [2016 – 2017].
2. [Journal Editorship] Associate Editor for IEEE Transactions on Knowledge and Data Engineering (TKDE) [2016 – Present]; Guest Associate Editor for MIS Quarterly, EPJ Data Science, Intl Journal of Computer Networks, etc.; Member of the Review Board for Proceedings of the VLDB Endowment (PVLDB) [2016-2017].
3. [Conference Organization] Program Committee Co-Chair for IEEE International Conference on Intelligence and Security Informatics (ISI) [2020]; Finance Chair [2018] and Mentorship Co-Chair [2012] for ACM SIGMOD International Conference on Management of Data (SIGMOD); Program Committee Co-Chair for The IEEE Symposium on Privacy-Aware Computing (PAC) [2018]; Program Committee Co-Chair for The International Conference on Web-Age Information Management (WAIM) [2016]; Poster & Demo Co-Chair for ACM Conference on Computer and Communications Security (CCS) [2012]; Program Committee Members for more than 100 conferences (e.g., SIGMOD, KDD, ICDE, CIKM).
4. [Grant Review] National Science Foundation [2007-2009, 2011-2015, 2017-2020]; Army Research Office [2016-2017]; Defense Advanced Research Projects Agency [2016].

### *Teaching and Mentoring*

1. [Analytics Courses] Analytics Practicum I & II (AU), Web Analytics (Penn State & GW), Database Systems (GW & UTA), Data Mining (UTA), Social/Legal/Ethical Impact of Computing (GW),
2. [Cybersecurity Courses] Cybersecurity Analytics (AU), Data Privacy and Security (Penn State, GW, & UTA), Network Security (GW), Information Security II (UTA)
3. [Information Systems Courses] Edge of Information Technology (AU), Computer Networks (GW)
4. [Ph.D. & M.S. Supervision] Supervised 8 Ph.D. Dissertations and 2 M.S. theses. Placement after graduation: Google Research, Uber, Oracle, Amazon, Intel, J.P. Morgan, JD Research, etc.