
PROFESSIONAL EXPERIENCE	Citadel LLC Senior Data Scientist II , Data Strategies Group Data Scientist , Data Strategies Group	Chicago, IL, USA 2021 – <i>present</i> 2017 – 2021
	<ul style="list-style-type: none">• Managed the data featurization team and developed alpha-generating features from trillion-scale noisy data sets. Research outputs (including forecasts, machine learning algorithms, and web apps) directly impact trading decisions of investment teams in equities, fixed income, commodities, and more.• Pioneered extensive research on the firm’s most important alternative data sets, and became the firm’s widely-acknowledged expert on representative panel construction, bias correction, cohort comparison, and many more. Keynote speaker at Citadel’s annual Alternative Data Summit since 2018.• Built the alternative data observatory which provides a real-time and comprehensive read on economic activities, became the one-stop-shop for alternative data signals, and is used weekly by >30% members of the Citadel Investment Group.	
	University of Notre Dame Research Assistant , Center for Network & Data Science (iCeNSA)	Notre Dame, IN, USA 2012 – 2017
	<ul style="list-style-type: none">• Designed the higher-order network construction algorithm, and demonstrated real-world applications on clustering, ranking, and anomaly detection. Extensive research outputs including journal papers in the high impact interdisciplinary journals AAAS Science Advances, Scientific Reports, EPJ Data Science, tutorial at ACM SIGKDD, and workshop organizer at NetSci.• Designed a data-driven approach to inform the management of invasive species through global shipping using multiple data sources in close collaboration with ecologists and policy experts, with publications on ACM SIGKDD and PLOS ONE.• Modeled information diffusion across a large network of retail traders by tracking retweets of breaking news, and quantitatively evaluated the influence of social network dynamics on retail trading, resulting in a finance paper and a U.S. provisional patent.	
	U.S. Army Research Lab Research Intern , Network Science-CTA	Adelphi, MD, USA 2015 & 2016
	<ul style="list-style-type: none">• Designed the workflow of HONVis, the state-of-the-art interactive visualization software of higher-order networks, and published the design on IEEE PacificVis.	
	IBM Research Research Intern , Smarter Urban Dynamics Group	Dublin, Ireland 2014
	<ul style="list-style-type: none">• Analyzed temporal patterns of 200k users’ cellular data usage records, and developed proactive customer care strategies for telecom companies.	
	Fudan University Research Assistant , Adaptive Networks and Control Lab	Shanghai, China 2009 – 2012
	<ul style="list-style-type: none">• Revealed human interaction patterns in different contexts, resulting in a journal paper in IEEE Transactions on Systems, Man, and Cybernetics.	
EDUCATION	University of Notre Dame Ph.D., Computer Science and Engineering, GPA 4.0/4.0 Dissertation: <i>Representing Big Data as Networks: New Methods and Insights</i> Advisor: Dr. Nitesh V. Chawla	Notre Dame, IN, USA 2017
	Fudan University B. Sc., Electronics Engineering	Shanghai, China 2012

RESEARCH INTERESTS	<p>Data Mining and Network Science</p> <p>Applications: using data science and network models to solve interdisciplinary problems in complex systems such as financial market (e.g., information diffusion & trading behavior), social network (e.g., online / mobile phone social interactions), and biology (e.g., species invasions via global shipping) in close collaboration with domain experts.</p> <p>Theory: network representation of various types of data; in particular, methods and influences of embedding rich information such as higher-order dependencies into networks.</p>	
TECHNICAL SKILLS	<p>Programming: Python and SQL for data mining; C# for graphical interface; C and Common Lisp for high performance computing.</p> <p>Big data: Snowflake, BigQuery, Vertica, and Spark for trillion-scale database queries; Airflow and Linux shell scripts for managing parallelized jobs on Amazon S3 and distributed systems.</p> <p>Tools: Tableau and Gephi for visualization; Bloomberg Terminal for trading; ArcGIS for geographic information system; NetworkX for network analysis; Javascript and Flask for web app development; Requests and Selenium for web scraping.</p>	
AWARDS AND HONORS	<p>Advanced Teaching Scholars Certificate, Notre Dame</p> <p>Outstanding Research Poster Award – Faculty Vote, Notre Dame</p> <p>1st Prize as team leader, Schurz Innovation Award on Data Mining, Notre Dame</p> <p>Outstanding Bachelor Thesis Award, Fudan University</p> <p>National Outstanding College Student (10 students per year), Chinese Ministry of Education</p> <p>2nd Prize as team leader, National Undergraduate Electronic Design Contest, Shanghai, China</p> <p>Outstanding Student, Fudan University (top 5% in the university)</p> <p>1st Prize, National Olympiad in Informatics in Provinces (NOIP), Hubei, China</p>	<p>2017</p> <p>2015</p> <p>2013</p> <p>2012</p> <p>2011</p> <p>2011</p> <p>2010</p> <p>2006</p>
PUBLICATIONS	<ol style="list-style-type: none"> 1. J. Xu, T. L. Wickramaratne, and N. V. Chawla. Representing higher-order dependencies in networks. <i>Science Advances</i>, 2(5):e1600028, 2016. (In the top 5% of all research outputs scored by Altmetric, 97th percentile attention score compared to outputs of the same age, featured in 8 news outlets. Journal impact factor 11.5) 2. M. Saebi, J. Xu, L. M. Kaplan, B. Ribeiro, and N. V. Chawla. Efficient modeling of higher-order dependencies in networks: from algorithm to application for anomaly detection. <i>EPJ Data Science</i>, 9(1):1–22, 2020 3. J. Xu, T. L. Wickramaratne, N. V. Chawla, E. K. Grey, K. Steinhäuser, R. P. Keller, J. M. Drake, and D. M. Lodge. Improving management of aquatic invasions by integrating shipping network, ecological, and environmental data: Data mining for social good. In <i>Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining (KDD)</i>, pages 1699–1708. ACM, 2014. (Acc. rate: 15%) 4. Y. Dong, R. A. Johnson, J. Xu, and N. V. Chawla. Structural diversity and homophily: a study across more than one hundred big networks. In <i>Proceedings of the 23rd ACM SIGKDD international conference on Knowledge discovery and data mining (KDD)</i>. ACM, 2017. (Acc. rate: 17%). 5. M. Saebi, J. Xu, S. R. Curasi, E. K. Grey, N. V. Chawla, and D. M. Lodge. Network analysis of ballast-mediated species transfer reveals important introduction and dispersal patterns in the arctic. <i>Scientific reports</i>, 10(1):1–15, 2020 6. M. Saebi, J. Xu, E. K. Grey, D. M. Lodge, J. J. Corbett, and N. Chawla. Higher-order patterns of aquatic species spread through the global shipping network. <i>PLOS ONE</i>, 	

15(7):e0220353, 2020

7. J. **Xu** and N. V. Chawla. Mining features associated with effective tweets. In *The international conference on Advances in Social Network Analysis and Mining (ASONAM)*. IEEE/ACM, 2017
8. J. Tao, J. **Xu**, C. Wang, and N. V. Chawla. HoNVis: Visualizing and exploring higher-order networks. In *IEEE Pacific Visualization Symposium (Pacific Vis)*. IEEE, 2017.
9. J. **Xu**, J. Tao, N. V. Chawla, and C. Wang. Visual analytics of higher-order dependencies in sensor data: Demo abstract. In *Proceedings of the Second International Conference on Internet-of-Things Design and Implementation*, pages 297–298. ACM, 2017.
10. Y.-Q. Zhang, X. Li, J. **Xu**, and A. V. Vasilakos. Human interactive patterns in temporal networks. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 45(2):214–222, 2015. (Journal impact factor 5.1%).
11. X. Wu, Y. Dong, H. Chao, J. **Xu**, D. Wang, and N. V. Chawla. Uapd: Predicting urban anomalies from spatial-temporal data. In *European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD)*. ECML-PKDD, 2017

SOFTWARE PACKAGES

1. HON: efficient construction and benchmarking of the higher-order network. Available at <https://github.com/xyjprc/hon>
Video demo available at <https://www.HigherOrderNetwork.com>
2. HoNVis: visualization software for HON. Developed with Dr. Jun Tao. Available at <https://www.HigherOrderNetwork.com>
3. Method and Apparatus for Effective Communications. U.S. Provisional Patent, number 61/855,060, filed May 7, 2013
4. TemporalMotiffinder: searching for temporal motifs in temporal networks. Software copyright of China, Jun. 2012
5. Motiffinder: efficiently searching for motifs in complex networks. Software copyright of China, 2011SR027720, May 2011

PRESENTATIONS, SEMINARS, AND POSTERS

1. Alternative data landscape. *Invited talk as guest lecturer for the class “Portfolio Management” and the class “Financial Innovation”*, UIUC, 2021 and 2022
2. Alternative data for fixed income. *Invited talk as guest lecturer for the class “Fixed Income Securities”*, Notre Dame, 2022
3. Higher-order networks for anomaly detection. *2nd ACM SIGKDD Workshop on Anomaly Detection in Finance (KDD-ADF)*, Anchorage, AK, Aug. 2019
4. Visual exploration of social signals with HON. *International School and Conference on Network Science (NetSci) Satellite: Higher-order Models in Network Science*, Indianapolis, IN, Jun. 2017
5. Visual exploration of social signals with HON. *U.S. Army Research Lab*, Aberdeen Proving Ground, MD, Mar. 2017
6. Representing higher-order dependencies in networks. *Northwestern Institute on Complex Systems (NICO)*, Northwestern University, Evanston, IL, Feb. 2017
7. Representing higher-order dependencies in networks. *Center for Complex Network Research (CCNR)*, Northeastern University, Boston, MA, Dec. 2016
8. From complex interactions to networks: capturing crucial patterns with the higher-order network. *U.S. Army Research Lab*, Adelphi, MD, Aug. 2016
9. Representing higher-order dependencies in networks. *International School and Conference on Network Science (NetSci)*, Seoul, Korea, Jun. 2016

10. Structural diversity and homophily: A study across more than one hundred large-scale networks. *2nd Annual International Conference on Computational Social Science (IC2S2)*, Chicago, Jun. 2016
11. Representing higher-order dependencies in networks. *2nd Annual International Conference on Computational Social Science (IC2S2)*, Chicago, Jun. 2016
12. Representing higher order dependencies in networks, and how it improves the result of a variety of tasks such as random walking, clustering, ranking, and anomaly detection. *Data, Algorithms and Problems on Graphs (DAPG)*, New York, Sep. 2015
13. Catching fire: Impact of information diffusion via twitter on trading, return, and liquidity. *TD Ameritrade*, New York, NY, Jun. 2015
14. Nonindigenous species risk assessment and prediction system (NIS-RAPS). *National Science Foundation Annual Report*, Feb. 2015
15. How data mining helps solve aquatic invasions for social good? *20th ACM SIGKDD international conference on Knowledge discovery and data mining*, Aug. 2014
16. Profiling human mobility based on digital traces. *IBM Research Dublin*, Ireland, Aug. 2014
17. Strategies for effective tweeting. *Schurz Innovation Award on Data Mining*, Notre Dame, May. 2013 (**Featured on South Bend Tribune and Notre Dame website**)

PROFESSIONAL SERVICES

Conference organization

- International School and Conference on Network Science Satellite: Higher-order Models in Networks (HONS @ NetSci 2018, Paris, France; as organizer & session chair)

Tutorial

- ACM SIGKDD all-day hands-on tutorial: Beyond Graph Mining: Higher-Order Data Analytics for Temporal Network Data (KDD 2018, London, UK)

Program committee member

- SIAM Data Mining Conferences (SDM 2024 @ Houston, SDM 2023 @ Minneapolis, SDM 2022 @ Alexandria, SDM 2021 @ Alexandria, SDM 2020 @ Cincinnati, SDM 2017 @ Houston)
- The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML-PKDD 2022 @ Grenoble, France, ECML-PKDD 2020 @ Ghent, Belgium)
- KDD Workshop on Machine Learning in Finance (KDD-MLF 2023 @ Long Beach, KDD-MLF 2022 @ Washington D.C., KDD-MLF 2021 @ Virtual, 2020 @ San Diego, USA)
- NIPS 2019 Workshop on Robust AI in Financial Services (Robust AI @ NIPS 2019, Vancouver, Canada)
- 2nd ACM SIGKDD Workshop on Anomaly Detection in Finance (KDD-ADF 2019 @ Anchorage, Alaska, USA)
- Big Network Analytics Workshop at World Wide Web Conference (BigNet @ WWW 2017, Perth, Western Australia)

External reviewer

- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- IEEE Transactions on Systems, Man and Cybernetics: Systems (SMC)
- IEEE Transactions on Big Data
- IEEE Transactions on Neural Networks and Learning Systems (TNNLS)
- Big Data Journal
- Knowledge and Information Systems (KAIS)
- Social Network Analysis and Mining (SNAM)
- Science of the Total Environment (STOTEN)

- Ocean and Coastal Management
- AAAI Conference on Artificial Intelligence (AAAI), 2016
- Advances in Social Network Analysis and Mining (ASONAM), 2013

TEACHING

- Implemented a new teaching strategy for the graduate-level Machine Learning course in Spring 2017 and completed a teaching experience mentored by Dr. Nitesh Chawla
- Advanced Teaching Scholar Certificate, Notre Dame, Jun. 2017 (as the second graduate student in Computer Science that ever received the award since the program's inception in 2007)
- Striving for Excellence in Teaching Certificate, Notre Dame, Nov. 2016

OTHER PASSIONS

- I enjoy doing random walks on the Wikipedia network (for hours) and learn about different aspects of our world.
- I am passionate about knowledge sharing; videos I made with college roommates on electronics DIY had >1M views back in the 2000s, making us the first-generation online influencers in China.
- I love making tools and optimizing daily routines; I built a smart radio that reads me customized financial news in the morning, and how well my crypto trading script had done overnight.