Xiaoran Yan

Curriculum Vitae
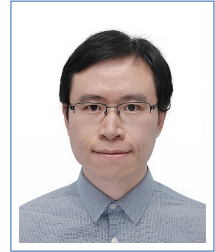
1001 E SR 45/46 Bypass
Bloomington, IN 47408
✆ (505) 554 4266
☎ (812) 856 4159
✉ xiaoran.a.yan@gmail.com
🖱 https://xiaoranyan.wordpress.com/

Research interests

Network Science (methods, theories, applications), Graph Theory, Machine Learning, Data Mining, Statistical Physics, Bioinformatics, Computational Neuroscience, Computational Social Science, Science of Science, Scientific Reproducibility

Education

2007–2013 **Ph.D. Computer Science**, *University of New Mexico*, USA, Advisor: Cristopher Moore.
Dissertation: *Model Selection for Stochastic Block Models.*

2003–2007 **B.S. Computer Science**, *Zhejiang University*, China, Advisor: Xiaogang Jin.
Thesis: *Dynamics of Cellular Automata on Small-World Networks.*

Employment

2015–Present **Assistant Research Scientist** , *Network Science Institute,* Indiana University.

2013–2015 **Postdoctoral Research Associate**, *Information Sciences Institute,* University of Southern California, Supervisor: Kristina Lerman.

2012–2013 **Graduate Fellow**, Santa Fe Institute, Supervisor: Cristopher Moore.

2009–2012 **Research Assistant**, *Computer Science Department,* University of New Mexico, Supervisor: Cristopher Moore.

2012 & 2007–2009 **Teaching Assistant**, *Computer Science Department,* University of New Mexico.

Honors and Awards

2017 **Travel Subsidy Award**, *Arthur M. Sackler Colloquia on Reproducibility of Research*, National Academy of Sciences, USA.

2016 **Travel Subsidy Award**, *Big Data Neuroscience Workshop*, Advanced Computational Neuroscience Network, USA.

2012–2013 **Graduate Fellowship**, *Santa Fe Institute*, NM, USA.

2007 **Outstanding B.S. Thesis Award**, *Zhejiang University*, China.

2003 **Zhao Zhongsu Award for Best Ningbo High Students**, China.

Research Grants and Support

2019-2022(2024) **Science Genome: A Scholarly Graph Embedding Framework to Uncover the Fundamental Dynamics of Scientific Enterprise**, *Air Force Office of Scientific Research, Minerva Research Initiative*, $4,394,586 total for 5 years, including the final two pending years.
Role: Senior personnel (10% FTE for 5 years), PI: Yong Yeol Ahn

2018-2020 **Shared Big Data Gateway: A Cloud-based infrastructure for Sharing Research Assets and Advancing Library and Information Science**, *Institute of Museum and Library Services. NLG-Libraries-FY18-2, National Leadership Grants for Libraries*, $2,126,137 total for 2 years, including cost share.
Role: Co-project Director (20% FTE for 2 years), PD: Wittenberg, J., Other Co-PDs: Pentchev, V., Mabry, P.L., Van Rennes, R

2018-2019 **Imaging and Genetic Biomarkers of Alzheimer's Disease and Imaging Epigenetics of Alzheimer's Disease**, *NIH R01AG040770-05 ($1,374,099 for 5 years) and NIH/NIA K02AG048240 ($434,160 for 3 years)*.
Role: Senior personnel (33% FTE for 6 months), PI: Apostolova, L.

2018-2019 **Comparing Microsoft Academic Graph with Web of Science** , *Microsoft Research*, Azure Research Award, $10,000 in terms of Azure credits over 1 year.
Role: PI

2017 **Strengthening Reproducibility in Network Science, Workshop held at NetSci2017**, *Microsoft Research*, Workshop Sponsorship, $6,000.
Role: Co-organizer, Other Co-organizers: Wang, K., Mabry, P.L., Fortunato, S.

2013-2015 **Inferring Structure and Forecasting Dynamics on Evolving Networks**, *AFOSR-MURI Grant No. FA955-10-1-0569*, $1,395,694.
Role: Postdoctoral Research Associate, PI Jeffrey Brantingham, Co-PI: Kristina Lerman

2013-2015 **Algebraic Graph Algorithms: the Laplacian and Beyond**, *NSF Award No. CCF-1111270*, $724,700.
Role: Research Collaborator, PI: Shanghua Teng

2013-2015 **Rethinking Network Structure: the Role of Interactions in the Analysis of Network Structure**, *NSF Award No. CIF-1217605*, $450,000.
Role: Postdoctoral Research Associate, PI: Kristina Lerman

2012-2013 **Statistical Inference for Detecting Structures and Anomalies in Networks**, *AFOSR-DARPA Grant No. FA9550-12-1-0432*, $2,924,396.
Role: Graduate Fellow, PI: Cristopher Moore

2009-2011 **Statistical Inference and Machine Learning for Complex Networks**, *James S. McDonnell Foundation Grant No. 220020171*, $417,576.
Role: Research Assistant, PI: Cristopher Moore

Peer-reviewed publications

Yu S., Xia F., Sun Y., Tang T., **Yan X.**, Lee I.(2020). Detecting Outlier Patterns with Query-based Artificially Generated Searching Conditions. To appear in the IEEE Transactions on Computational Social Systems

Yan, X., Sporns, O., and Avena-Koenigsberger, A. (2020). Efficient network navigation with partial information. To appear in the Proceedings of the 2020 IEEE International Conference on Networking, Sensing and Control (ICNSC2020)

Kong, X., Xia, F., Fu, Z., **Yan, X.**, Tolba, A., and Almakhadmeh, Z. (2019). TBI2Flow: Travel behavioral inertia based long-term taxi passenger flow prediction. *World Wide Web Journal*. Special Issue on Smart Computing and Cyber Technology for Cyberization

Ploszaj, A., **Yan, X.**, and Borner, K. (2019). The impact of air transport availability on research collaboration. In Proceedings of the ISSI 2019 Conference

Avena-Koenigsberger, A., **Yan, X.**, Kolchinsky, A., Van den Heuvel, M., Hagmann, P., and Sporns, O. (2019). A spectrum of routing strategies for brain networks. *PLOS Computational Biology*. 15(3), pp.e1006833-e1006833

Yan, X., Jeub, L., Flammini, A., Radicchi, F., and Fortunato, S. (2018b). Weight Thresholding on Complex Networks. *Phys. Rev. E*, 98:042304

Faskowitz, J., **Yan, X.**, Zuo, X.-N., and Sporns, O. (2018). Weighted stochastic block models of the human connectome across the life span. *Scientific reports*, 8(1):12997

Yan, X., Sadler, B. M., Drost, R. J., Yu, P. L., and Lerman, K. (2017a). Graph Filters and the Z-Laplacian. *IEEE Journal of Selected Topics in Signal Processing*, 11(6):774–784

Yan, X., Teng, S.-H., and Lerman, K. (2017b). Multi-layer Network Composition Under a Unified Dynamical Process. In *Social, Cultural, and Behavioral Modeling: 10th International Conference, SBP-BRiMS 2017, Washington, DC, USA, July 5-8, 2017, Proceedings*, pages 315–321. Springer International Publishing. DOI: 10.1007/978-3-319-60240-0_38

Merkurjev, E., Bertozzi, A., **Yan, X.**, and Lerman, K. (2017). Modified Cheeger and ratio cut methods using the Ginzburg–Landau functional for classification of high-dimensional data. *Inverse Problems*, 33(7):074003

Yan, X. (2016). Bayesian model selection of stochastic block models. In *2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, pages 323–328

Yan, X., Teng, S.-h., Lerman, K., and Ghosh, R. (2016). Capturing the interplay of dynamics and networks through parameterizations of Laplacian operators. *PeerJ Computer Science*, 2:e57

Lerman, K., **Yan, X.**, and Wu, X.-Z. (2016). The "Majority Illusion" in Social Networks. *PLoS ONE*, 11(2):1–13

Daianu, M., Ver Steeg, G., Mezher, A., Jahanshad, N., Nir, T. M., **Yan, X.**, Prasad, G., Lerman, K., Galstyan, A., and Thompson, P. M. (2016). Information-Theoretic Clustering of Neuroimaging Metrics Related to Cognitive Decline in the Elderly. In Menze, B., Langs, G., Montillo, A., Kelm, M., Müller, H., Zhang, S., Cai, W., and Metaxas, D., editors, *Medical Computer Vision: Algorithms for*

Big Data: International Workshop, MCV 2015, Held in Conjunction with MICCAI 2015, Munich, Germany, October 9, 2015, Revised Selected Papers, pages 13–23. Springer International Publishing

Gupta, S., **Yan, X.**, and Lerman, K. (2015). Structural Properties of Ego Networks. In *Social Computing, Behavioral-Cultural Modeling, and Prediction*, volume 9021 of *Lecture Notes in Computer Science*, pages 55–64. Springer International Publishing

Ghosh, R., Teng, S.-h., Lerman, K., and **Yan, X.*** (2014). The Interplay Between Dynamics and Networks: Centrality, Communities, and Cheeger Inequality. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '14, page 1406–1415, New York, NY, USA. ACM, ACM

Yan, X., Jensen, J. E., Krzakala, F., Moore, C., Shalizi, C. R., Zdeborová, L., Zhang, P., and Zhu, Y. (2014). Model selection for degree-corrected block models. *Journal of Statistical Mechanics: Theory and Experiment*, 2014(5):P05007

Zhu, Y., **Yan, X.**, and Moore, C. (2014). Oriented and degree-generated block models: generating and inferring communities with inhomogeneous degree distributions. *Journal of Complex Networks*, 2(1):1–18

Zhu, Y., **Yan, X.**, Getoor, L., and Moore, C. (2013). Scalable Text and Link Analysis with Mixed-topic Link Models. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '13, page 473–481, New York, NY, USA. ACM, ACM

Moore, C., **Yan, X.***, Zhu, Y., Rouquier, J.-B., and Lane, T. (2011). Active Learning for Node Classification in Assortative and Disassortative Networks. In *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '11, page 841–849, New York, NY, USA. ACM, ACM

▆▆▆▆ Other publications

Yan, X., Avena-Koenigsberger, A., and Sporns, O. (2018a). Efficient routing with partial information. GTA3 2018: Workshop on Graph Techniques for Adversarial Activity Analytics

Wood, I., **Yan, X.**, Liu, X., and Ahn, Y.Y. (2017). Community Detection with Selective Zooming. Netsci 2017

Yan, X., Krzakala, F., Moore, C., Shalizi, C. R., Zdeborová, L., and Zhu, Y. (2012). Model selection for degree-corrected block models. NIPS Workshop on Social Network and Social Media Analysis: Methods, Models and Applications

Yan, X., Zhu, Y., Rouquier, J.-B., and Moore, C. (2009). Active Learning for Hidden Attributes in Networks. NIPS Workshop on Analyzing Networks and Learning with Graphs

Invited talks

2019, Jul **CADRE: Building a reproducible big data infrastructure for academic data**, *The 3rd Forum of Young Scholars*, Nanjing University, Nanjing, China, Oral presentation.

2018, Dec **Collaborative Archive & Data Research Environment (CADRE): Building a community infrastructure for big academic data**, *Alpha Lab*, Dalian University of Technology, Dalian, China, Oral presentation.

2018, Jul **Promoting openness and reproducibility in science of science research**, *School of Computer Science and Technology*, University of Science and Technology, Hefei, China, Oral presentation.

2017, Apr **Z-Laplacian: a dynamical framework for community structures and centralities on networks**, *American Mathematical Society Spring Central Section Meeting*, Indiana University, Oral presentation.

2017, Mar **Centralities, Communities and Dynamical processes on networks: an intuitive framework**, *School of Information Sciences*, University of Pittsburgh, Big Data Colloquium.

2016, Dec **Centralities, Communities and Dynamical Processes on Networks: The Z-Laplacian Framework**, *School of Computer Science and Technology*, University of Science and Technology of China, Colloquium.

2016, Oct **Correlating Air Transportation with Co-affiliation and Collaboration Data**, *Network Science Institute*, Indiana University, Open Science Meetings.

2016, Apr **Dynamical processes on networks and graph transformations**, *School of Informatics and Computing*, Indiana University, Intelligent & Interactive Systems Talk Series.

2015, Sep **Dynamical Processes and Graph Transformations**, *Senseable City Laboratory*, Massachusetts Institute of Technology, Pecha Kucha Seminar.

2015, Mar **Centralities, Communities and Dynamics: The Generalized Laplacian Framework**, *Imaging Genetics Center, Laboratory of Neuro Imaging*, Keck School of Medicine, University of Southern California, Seminar.

2014, May **Centralities, Communities and Operators: A Framework for Network Dynamics**, *Santa Fe Institute*, Seminar.

2013, Aug **Variational Inference of Community Models: A Unifying Learning Framework**, *Information Sciences Institute*, Viterbi School of Engineering, University of Southern California, AI Seminar.

2013, Jun **Scalable Model Selection for Networks using Belief Propagation**, *Technical University of Denmark*, Netsci 2013, Satellite symposium: Complex Networks meet Machine Learning.

2013, May **Model Selection for Stochastic Block Models**, *Santa Fe Institute*, Workshop on Structure, Statistical Inference and Dynamics in Networks: From Graphs to Rich Data.

Conference activities

2019, Sep **CADRE: A Platform for Enabling Research via Shared Data, Resources, and Community**, *17th International Conference on Scientometrics & Informetrics*, Sapienza University, Rome, Italy, Tutorial host and demo presenter.

2019, Jul **Evaluating the Utility of Homological Structural MRI Features and A Kernel Based Learning Framework for Prediction of MCI**, *The Alzheimer's Association International Conference 2019*, Los Angeles, CA, US, Poster presentation.

2018, Nov **The Shared Big Data Gateway (SBD-Gateway) project: Collaborative Archive & Data Research Environment (CADRE)**, *All Hands Meeting of the Midwest Big Data Hub*, Columbus, Ohio, US, Poster presentation.

2018, Feb **Efficient routing with partial information**, *GTA3 2018: Workshop on Graph Techniques for Adversarial Activity Analytics*, Marina del Rey, CA, US, Oral presentation.

2017, Jul **Multi-layer Network Composition under a Unified Dynamical Process**, *SBP and BRiMS 2017 Conference*, Washington DC, US, Oral presentation.

2017, Jun **Efficient Thresholding of Weighted Networks**, *NetSci 2017*, Indianapolis, IN, US, Oral presentation.

2016, Aug **Bayesian Model Selection of Stochastic Block Models**, *ASONAM 2016*, San Francisco, CA, US, Oral presentation.

2016, May **Graph Filters and the generalized Laplacian**, *Graph Signal Processing Workshop*, Philadelphia, PA, US, Oral presentation.

2015, Apr **Structural Properties of Ego Networks**, *SBP and BRiMS 2015 Conference*, Washington DC, US, Oral presentation.

2014, Aug **The Interplay Between Dynamics and Networks: Centrality, Communities, and Cheeger Inequality**, *ACM SIGKDD 2014 Conference*, New York, NY, US, Poster presentation.

2013, Aug **Scalable Text and Link Analysis with Mixed-topic Link Models**, *ACM SIGKDD 2013 Conference*, Chicago, IL, US, Poster presentation.

2012, Dec **Model selection for degree-corrected block models**, *NIPS 2012 Workshop on Social Networks and Social Media Analysis: Methods, Models and Applications*, Lake Tahoe, NV, US, Poster presentation.

2012, Jun **Oriented and degree-generated block models**, *NetSci 2012 Conference*, Evanston, IL, US, Poster presentation.

2011, Aug **Active Learning for Node Classification in Assortative and Disassortative Networks**, *ACM SIGKDD 2011 Conference*, San Diego, CA, US, Poster presentation.

2009, Dec **Active Learning for Hidden Attributes in Networks**, *NIPS 2009 Workshop on Analyzing Networks and Learning with Graphs*, Whistler, BC, Canada, Poster presentation.

Other professional activities

Co-organizer of the workshop and tutorial of "CADRE: A Platform for Enabling Research via Shared Data, Resources, and Community" at ISSI 2019

Co-organizer of the Satellite "Strengthening Reproducibility in Network Science" at NetSci 2017

Served in program committees of International AAAI Conference on Web and Social Media 2016, 2018, The Florida Artificial Intelligence Research Society 2016, 2017, NetSci 2017, SBP-BRiMS 2017, 2018, Complex Networks 2019, 2020, ACM-WSDM 2019, WWW 2018, 2019, 2020

Peer review of manuscripts for Proceedings of the National Academy of Sciences, Journal of Complex Networks, Physica A, IEEE Transactions on Signal Processing, Computer Science Review, ACM SIGKDD 2014, Socinfo 2014, WWW 2015, NIPS 2016, ICLR 2018

Professional member of ACM, IEEE, ASSS and ISSI. Member of the Open Acadmic Society.

Teaching

2019 **Bi-weekly reading group on graph embedding algorithms**, *Indiana University*.
Organizer

2019 **Mentoring an undergraduate student (Xu, F.) for his thesis**, *University of Heilongjiang*.
Mentor

2016 **Weekly interdisciplinary reading group including students from computer science, informatics and neuroscience**, *Indiana University*.
Organizer

2014 **Mentoring an undergraduate student (Gupta, S.) under the supervision of Kristina Lerman**, *University of Southern California*.
Mentor

2012 **Introduction to the Theory of Computation**, *University of New Mexico*, Graduate level course.
Teaching Assistant and Guest Lecturer

2009 **Lab sessions of Computer Programming Fundamentals**, *University of New Mexico*, Undergraduate level course.
Instructor

2008 **Design of Large Programs**, *University of New Mexico*, Undergraduate level course.
Teaching Assistant

2007 **Data Structures and Algorithms**, *University of New Mexico*, Graduate level course.
Teaching Assistant

Reference list

Andrea L. Bertozzi
Department of Mathematics
University of California Los Angeles
7619D Mathematical Sciences
Los Angeles, CA 90095
✉ bertozzi@math.ucla.edu
☎ 310-825-4340

Olaf Sporns
Department of Psychological and Brain
Sciences
Indiana University
1101 E. 10th Street
Bloomington, IN 47408
✉ osporns@indiana.edu
☎ 812-855-2772

Filippo Menczer
School of Informatics, Computing and
Engineering
Indiana University
901 E. 10th Street
Bloomington, IN 47408
✉ fil@indiana.edu
☎ 812-856-1377

Brian M. Sadler
Adelphi Laboratory Center
Army Research Laboratory
2800 Powder Mill Rd
Adelphi, MD 20783
✉ brian.m.sadler6.civ@mail.mil
☎ 301-394-3590

Kristina Lerman
Information Sciences Institute
University of Southern California
4676 Admiralty Way
Marina del Rey, CA 90292, USA
✉ kristina.lerman@gmail.com
☎ 310-448-8714

Paul M. Thompson
Institute for Neuroimaging and Informatics
University of Southern California
4676 Admiralty Way
Marina del Rey, CA 90292
✉ pthomp@usc.edu
☎ 323-442-7246

Lise Getoor
Department of Computer Science
University of California Santa Cruz
1156 High Street
Santa Cruz, CA 95064, USA
✉ getoor@soe.ucsc.edu
☎ 831-459-1489

Franco Pestilli
Department of Psychological and Brain
Sciences
Indiana University
1101 E. 10th Street
Bloomington, IN 47408
✉ franpest@indiana.edu
☎ 812-856-9967

Cristopher Moore
Santa Fe Institute
1399 Hyde Park Road
Santa Fe, NM 87501, USA
✉ moore@santafe.edu
☎ 505-984-8800

Katy Borner
Department of Intelligent Systems Engineering & Information and Library
Science
Indiana University
700 N. Woodlawn Ave
Bloomington, IN 47408
✉ katy@indiana.edu
☎ 812-856-7034

Santo Fortunato
School of Informatics, Computing and
Engineering
Indiana University
901 E. 10th Street
Bloomington, IN 47408
✉ santo@indiana.edu
☎ 812-856-5056