Yinghui Wu

CONTACT INFORMATION

Department of Computer and Data Sciences
Case School of Engineering
Case Western Reserve University
2101 Martin Luther King Jr. Dr
Cleveland, OH 44106

Office: Room 515, Olin Building
Phone: (216) 368-8829
E-mail: yxw1650@case.edu
Website: yinghwu.github.io

RESEARCH THEMES

- Data Management and Database Systems
- Graph Analytics and Network Science
- Data and Information Quality

EDUCATION

The University of Edinburgh, Edinburgh, United Kingdom · June 2011

Ph.D. in Computer Science (Advisor: Wenfei Fan)

Peking University, Beijing, China · June 2007

B.S. in Computer Science

Peking University, Beijing, China · June 2007

B.S. in Economics

PROFESSIONAL APPOINTMENTS

Case Western Reserve University, Cleveland, OH · *Apr.2022 - Present*

Theodore L. and Dana J. Schroeder
Endowed Associate Professor, *Department of Computer and Data Sciences*

Case Western Reserve University, Cleveland, OH · *Jun.2021 - Apr.2022*

Associate Professor, *Department of Computer and Data Sciences*

Case Western Reserve University, Cleveland, OH · Sep.2019 - Jun.2021

Assistant Professor, *Department of Computer and Data Sciences*

Pacific Northwest National Laboratory, Richland, WA · Sep.2017 - Present

Staff Scientist[1], *Advanced Computing, Mathematics, and Data Division*

Washington State University, Pullman, WA · Sep.2014 - Jun.2019

Assistant Professor, *School of Electrical Engineering and Computer Science*

University of California Santa Barbara, Santa Barbara, CA · Sep.2011 - Sep.2014

Research Scientist, *Department of Computer Science* (Mentor: Xifeng Yan)

---

[1]This is a joint appointment between CWRU and PNNL.

<table>
<tr><td>PROFESSIONAL
MEMBERSHIPS</td><td>Member, Association for Computing Machinery (ACM)</td></tr>
</table>

PROFESSIONAL
MEMBERSHIPS

Member, Association for Computing Machinery (ACM)

Member, ACM Special Interest Group on Management of Data (SIGMOD)

Member, Institute for Electrical and Electronics Engineers (IEEE)

Member, Network Science Collaborative Technology Alliance (NS-CTA)

AWARDS AND
HONORS

Best Reviewer Award (Research Track), IEEE ICDE Conference		2023
Service Excellence Award, Department of Computer & Data Sciences, CWRU		2021-2022
Best Paper Award, IEEE International Conf. on Big Data		2020
Distinguished Reviewer Award, International Conf. on VLDB		2019
Best Paper Candidate, IEEE ICDE Conference C.10 [2]		2019
ACM SIGMOD Research Highlight Award, ACM		2018
Best Paper Award, ACM SIGMOD Conference C.23		2017
Best Demonstration Award, International Conf. on Very Large Data Bases (VLDB) C.24		2017
Google Faculty Research Award [3], Google		2015
Best Paper Runner-Up, IEEE ICDE Conference C.37		2014
Best Poster Paper Award, IEEE International Conf. on Data Engineering (ICDE) C.39		2013

PUBLICATIONS

Citations: 3326; ***h-index 29***; ***i10-index 46*** *(as of May, 2023, source: Google Scholar)*

I publish in multiple fields with different standards for ordering authors.

- *'=': authors are listed in alphabetical order (equal contribution);*
- *First authors (or a major contributor if listed in alphabetical order): in **bold**;*
- *My students and research advisees: ⋆;*
- *Representative publications are marked with ✳.*

Refereed Journal Articles [4]

J.1 [‡] **Arman Ahmed, Sajan K.Sadanandan, Shikhar Pandey, Sagnik Basumallik, Anurag Srivastava, Yinghui Wu**. Event Analysis in Transmission Systems Using Spatial Temporal Graph Encoder Decoder (STGED). *IEEE Transactions on Power Systems*, 2022

J.2 Jianxin Li, Chengfei Liu, Ziyu Guan, Yinghui Wu, Special Issue on Decision Making in Heterogeneous Network Data Scenarios and Applications. *World Wide Web Journal*, 2022

J.3 [=] Sai Pushpak Nandanoori, Sheng Guan, Soumya Kundu, Seemita Pal, Khushbu Agarwal, Yinghui Wu, Sutanay Choudhury. Graph Neural Network and Koopman Models for Learning Networked Dynamics: A Comparative Study on Power Grid Transients Prediction, *IEEE Access*. 2022

---

[2] A pointer is provided to the corresponding papers in "Publications" for each publication award.

[3] This competitive program has an average award rate of 15%.

[4] Several Computer Science Journals accept articles that are extended from conference publications with significant and new contribution after thorough peer review. These articles are indicated with a ‡.

J.4   **Qi Song***, Peng Lin*, **Yinghui Wu**. Answering Why Questions for Attributed Graphs. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 2021

J.5   **Arman Ahmed***, Sajan Kaduvettykun, Anurag Srivastava, **Yinghui Wu**. Anomaly Detection, Localization and Classification using Drifting Synchrophasor Data Streams. *IEEE Transactions on Smart Grid*, 2021

J.6   **Peng Lin***, **Yinghui Wu**. Experience: Incompleteness Assessment and Ensemble-based Fact Prediction for Knowledge Graphs. *ACM Journal of Data and Information Quality (JDIQ)* (under revision), 2021

J.7[‡]   **Peng Lin***, Qi Song*, **Yinghui Wu**, Jiaxing Pi. Discovering Patterns for Fact Checking in Knowledge Graphs. *ACM Journal of Data and Information Quality (JDIQ)*, 11.3 (2019): 1-27.

J.8[‡]   **Xuanming Liu***, Tingjian Ge, Yinghui Wu. A Stochastic Approach to Finding Densest Temporal Subgraphs in Dynamic Graphs. *IEEE Transactions on Knowledge and Data Engineering (TKDE)* (accepted), 2020

J.9   **Arman Ahmad***, Vignesh V.G. Krishnan*, Seyedeh Foroutan*, Md. Touhiduzzaman, Caroline Rublein, **Anurag Srivastava**, **Yinghui Wu**, Adam Hahn, Sindhu Suresh. Cyber Physical Security Analytics for Anomalies in Transmission Protection Systems. *IEEE Transactions on Industry Applications*, vol. 55, no. 6, pp. 6313-6323, 2019

J.10   **Peng Lin***, Qi Song*, **Yinghui Wu**. Fact Checking in Knowledge Graphs with Ontological Subgraph Patterns. *Data Science & Engineering*, 3.4(2018): 341-358.

J.11[‡]   **Qi Song***, **Yinghui Wu**, Peng Lin*, Xin (Luna) Dong. Mining Summaries for Knowledge Graph Search. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 30(10), 1887-1900, 2018

J.12   **Mengze Zhou***, Yuhui Wang*, **Anurag K Srivastava**, **Yinghui Wu**, Paramarshi Banerjee. Ensemble Based Algorithm For Syncrophasor Data Anomaly Detection. *IEEE Transactions on Smart Grid*, 10.3 (2018): 2979-2988.

J.13   **Wenfei Fan**, **Yang Cao**, **Jingbo Xu**, **Wenyuan Yu**, **Yinghui Wu**, Chao Tian, Jiaxin Jiang, Bohan Zhang. From Think Parallel to Think Sequential. *SIGMOD Record*, 47.1: 15-22, 2018

J.14[=]   Wenfei Fan, Jingbo Xu, Xiaojian Luo, **Yinghui Wu**, Wenyuan Yu, and Ruiqi Xu. GRAPE: Conducting Parallel Graph Computations without Developing Parallel Algorithms. *IEEE Data Eng. Bull* 40(3): 30-41, 2017.

J.15[=]   Lawrence B.Holder, Rajmonda Caceres, David F.Gleich, Jason Riedy, Maleq Khan, Nitesh V.Chawla, Ravi Kumar, **Yinghui Wu**, Christine Klymko, Tina Eliassi-Rad, Aditya Prakash, Current and Future Challenges in Mining Large Networks: Report on the Second SDM Workshop on Mining Networks and Graphs. *ACM SIGKDD Explorations Newsletter* 18.1: 39-45, 2016.

J.16[=] [‡]Wenfei Fan, Xin Wang, **Yinghui Wu**. Answering graph query using views. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, vol. 28, no. 2, pp. 326-341. 2016.

J.17= ‡ Wenfei Fan, Xin Wang, **Yinghui Wu**. Incremental Graph Pattern Matching. *ACM Transactions on Database Systems (TODS)*, 38.3 (2013): 1-47.

J.18= Ting Deng, Wenfei Fan, Leonid Libkin, **Yinghui Wu**. On the Aggregation problem for Synthesized Web Services. *Journal of Computer and System Science (JCSS)*. 79.6 (2013): 873-891

J.19= ‡ Wenfei Fan, Jianzhong Li, Jizhou Luo, Zijing Tan, Xin Wang, **Yinghui Wu**. Adding regular expressions to graph reachability and pattern queries. *Frontiers of Computer Science*, 3.6 (2012): 313-338.

Refereed Conference Papers [5] [6]

C.1 Fotis Psallidas, Megan Eileen Leszczynski, Mohammad Hossein Namaki⋆, Avrilia Floratou, Ashvin Agrawal, Konstantinos Karanasos, Subru Krishnan, Pavle Subotic, Markus Weimer, Yinghui Wu, Yiwen Zhu. Demonstration of Geyser: Provenance Extraction and Applications over Data Science Scripts. *ACM SIGMOD Conference on Management of Data (SIGMOD),* 2023

C.2 Hanchao Ma⋆, Sheng Guan⋆, Mengying Wang⋆, Qi Song, Yinghui Wu. Fair Group Summarization with Graph Patterns. *Int.Conf.on Data Engineering (ICDE),* 2023

C.3 Yangxin Fan⋆, Xuanji Yu, Raymond Wieser, David Meakin, Avishai Shaton, Jean-Nicolas Jaubert, Robert Flottemesch, Michael Howell, Jennifer Braid, Laura Bruckman, Roger French, Yinghui Wu. Spatio-Temporal Denoising Graph Autoencoders with Data Augmentation for Photovoltaic Timeseries Data Imputation. *ACM SIGMOD Conference on Management of Data (SIGMOD),* 2023

C.4 Sheng Guan⋆, Hanchao Ma⋆, Mengying Wang⋆, Yinghui Wu. GALE: Active Adversarial Learning for Erroneous Node Detection in Graphs. *Int.Conf.on Data Engineering (ICDE),* 2023

C.5 Morteza Alipour Langouri⋆, Adam Mansfield, Fei Chiang, Yinghui Wu. Inconsistency Detection with Temporal Graph Functional Dependencies. *Int.Conf.on Data Engineering (ICDE),* 2023

C.6 Mengying Wang⋆, Hanchao Ma⋆, Abhishek Daundkar⋆, Sheng Guan⋆, Yiyang Bian⋆, Alp Sehirlioglu, Yinghui Wu. CRUX: Crowdsourced Materials Science Resource and Workflow Exploration. *The 31st ACM International Conference on Information and Knowledge Management (CIKM)*, 2022

C.7 Liangyi Huang⋆, Sophia R.Hall, Fei Shao, Arafath Nihar, Vipin Chaudhary, Yinghui Wu, Roger French, Xusheng Xiao. System-Auditing, Data Analysis and Characteristics of Cyber Attacks for Big Data Systems *The 31st ACM International Conference on Information and Knowledge Management (CIKM),* 2022

C.8 Sheng Guan⋆, Hanchao Ma⋆, Yinghui Wu. RoboGNN: Robustifying Node Classification under Link Perturbation, *The 31st International Joint Conference on Artificial Intelligence (IJCAI),* 2022

---

[5]Appeared in conferences that have high standards of novelty, quality and impact, and are considered equivalent to journal publications within the Computer Science community as well as by NSF guidelines.

[6]For each conference paper with peer-reviewed extended journal version, a pointer to the journal paper is provided.

C.9   Hanchao Ma⋆, Sheng Guan⋆, Mengying Wang⋆, Yen-shuo Chang⋆, Yinghui Wu. Subgraph Query Generation with Fairness and Diversity Constraints. *Int. Conf. on Data Engineering (ICDE),* 2022

C.10   Hanchao Ma⋆, Sheng Guan⋆, Christopher Toomey⋆, Yinghui Wu. Diversified Subgraph Query Generation with Group Fairness. *The 15th International Conference on Web Search and Data Mining (WSDM)*, 2022

C.11 William Oltjen⋆, Yangxin Fan⋆, Jiqi Liu⋆, Liangyi Huang⋆, Mingjie Li, Hubert Seigneur, Xusheng Xiao, Kristopher Davis, Laura Bruckman, Yinghui Wu, Roger French. FAIRification, Quality Assessment, and Missingness Pattern Discovery for Spatiotemporal Photovoltaic Data. *The 49th IEEE Photovoltaic Specialists Conference (PVSC)*, 2022

C.12   Sayak Mukherjee⋆, Sai Pushpack Nadanoori, Sheng Guan⋆, Khushubu Agarwal, Subhrajit Sinha, Soumya Kundu, Seemita Pal, Yinghui Wu, Draguna L Vrabie, Sutanay Choudhury. Learning Distributed Geometric Koopman Operator for Sparse Networked Dynamical Systems, *Learning on Graphs Conference (LoG).* 2022

C.13   **Sheng Guan**⋆, Hanchao Ma⋆, **Yinghui Wu**. GEDet: Detecting Erroneous Nodes with A Few Examples. *Very large Data Bases (VLDB)*, 2021.

C.14   **Qi Song**⋆, Hanchao Ma⋆, Peng Lin⋆, **Yinghui Wu**. GRIP: Constraint-based Explanation for Incomplete Graphs.*ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2021

C.15   **Qi Song**⋆, Hanchao Ma⋆, Peng Lin⋆, **Yinghui Wu**. Explaining Missing Data in Graphs: A Constraint-based Approach. *Int.Conf. on Data Engineering (ICDE)*, 2021.

C.16   **Ahmad Maroof Karimi**⋆, **Yinghui Wu**, Mehmet Koyuturk, Roger French. Spatialtemporal Graph Neural Network for Performance Prediction of Photovoltaic Power Systems. *Annual Conference on Innovative Applications of Artifical Intelligence (IAAI)*, 2021.

C.17   **Sheng Guan**⋆, Hanchao Ma⋆, **Yinghui Wu**. GEDet: Adversarially Learned Few-Shot Erroneous Node Detection in Graphs. *IEEE International Conference on Big Data (IEEE BigData)*, 2020.

C.18   **Mohammad Hossein Namaki**⋆, **Avrilia Floratou**, Fotis Psallidas, Subru Krishnan, Ashvin Agrawal, **Yinghui Wu**. Vamsa:Automated Provenance Tracking in Data Science Scripts. *Int. Conf. on Knowledge Discovery and Data Mining (KDD)*, 2020. (*16.8% acceptance rate*)

C.19   **Mohammad Hossein Namaki**⋆, Xin Zhang⋆, Sukhujinder Singh⋆, **Arman Ahmad**⋆, Armina Foroutan, **Yinghui Wu**, **Anurag Srivastava**, Anton Kocheturov. Kronos: Lightweight Knowledge-based Event Analysis in Cyber-Physical Data Streams. *Int.Conf.on Data Engineering (ICDE)*, 2020. (*Acceptance rate n/a; average acceptance rate (last 5 years):* 15.8%)

C.20* **Peng Lin**⋆, **Qi Song**⋆, **Yinghui Wu**, Jiaxing Pi. Repairing Entities using Star Constraints in Multirelational Graphs. *Int.Conf.on Data Engineering (ICDE)* , 2020. (*Acceptance rate n/a; average acceptance rate (last 5 years):* 15.8%)

C.21   **Sheng Guan**⋆, Hanchao Ma,⋆, **Yinghui Wu**. Attribute-Driven Backbone Discovery. *Int.*

Conf. on Knowledge Discovery and Data Mining (KDD), 2019. (*14.2% acceptance rate*)

C.22 **Hanchao Ma⋆**, **Morteza Aliour**, **Yinghui Wu**, Fei Chiang. Ontology-based Entity Matching in Attributed Graphs. *International Conference on Very Large Data Bases (VLDB)*, 2019. (*Acceptance rate n/a; average acceptance rate (last 5 years): 17.5%*)

C.23 **Mohammad Hossein Namaki⋆**, Qi Song⋆, **Yinghui Wu**. NAVIGATE: Explainable Visual Graph Exploration by Examples. *ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2019. (*20% acceptance rate*)

C.24* **Mohammad Hossein Namaki⋆**, Qi Song⋆, **Yinghui Wu**, Shengqi Yang. Answering Why-questions by Exemplars in Attributed Graphs. *ACM SIGMOD Conference on Management of Data (SIGMOD)* , 2019. (*20% acceptance rate*) [Journal Version: J.1]

C.25 **Qi Song⋆**, Mohammad Hossein Namaki⋆, **Yinghui Wu**.Answering Why-Questions for Subgraph Queries in Attributed Graphs.*International Conference on Data Engineering(ICDE)*, 2019. (*26.8% acceptance rate*)

C.26 **Xuanming Liu**, Tingjian Ge, **Yinghui Wu**. Finding Densest Lasting Subgraphs in Dynamic Graphs: a Stochastic Approach. *International Conference on Data Engineering (ICDE)* , 2019. (*26.8% acceptance rate*) (***Best Paper Candidate)*** [Journal Version: J.2]

C.27 **Mohammad Hossein Namaki⋆**, **Yinghui Wu**, Xin Zhang⋆. GExp: Cost-aware Graph Exploration with Keywords. *ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2018. (*20% acceptance rate*)

C.28 **Qi Song⋆**, **Bo Zong**, **Yinghui Wu**, Lu-An Tang, Hui Zhang, Guofei Jiang and Haifeng Chen. TGNet: Learning to Rank Nodes in Temporal Graphs *The 27th ACM International Conference on Information and Knowledge Management (CIKM)*, 2018 (*17.0% acceptance rate*)

C.29 **Peng Lin⋆**, Qi Song⋆, Jialiang Shen⋆, **Yinghui Wu**. Discovering Graph Patterns for Fact Checking in Knowledge Graphs. *23rd International Conference on Database Systems for Advanced Applications (DASFAA)*, 2018. (*Acceptance rate n/a; average acceptance rate (last 5 years): 18.4%*) [Journal Version: J.4]

C.30 **Keyvan Sasani⋆**, **Mohammad Hossein Namaki⋆**, **Yinghui Wu**, Assefaw Gebremedhin. Multi-metric Graph Query Performance Prediction. *23rd International Conference on Database Systems for Advanced Applications (DASFAA)*, 2018 (*Acceptance rate n/a; average acceptance rate (last 5 years): 18.4%*)

C.31 **Mohammad Hossein Namaki⋆**, **Yinghui Wu**, Xin Zhang⋆. Diversified Keyword Expansion on Multi-labeled Graphs. *Asia-Pacific Web (APWeb) and Web-Age Information Management (WAIM) Joint International Conference on Web and Big Data*, 2018 (*Acceptance rate n/a*)

C.32 **Sutanay Choudhury**, **Sumit Purohit**, **Peng Lin⋆**, **Yinghui Wu**, Lawrence Holder, Khushbu Agarwal. Percolator: Scalable Pattern Discovery in Dynamic Graphs. *ACM International Conference on Web Search and Data Mining (WSDM)*, 2018 (*16.1% acceptance rate*)

C.33 **Arman Ahmad⋆**, **Vignesh V.G. Krishnan⋆**, **Anurag Srivastava**, **Yinghui Wu**, Suresh

Sindhu. Cyber Physical Security Analytics for Transactive Energy Systems Using Ensemble Machine Learning. *North American Power Symposium (NAPS)*, 2018 (*Acceptance rate n/a*)

C.34 **Arman Ahmad⋆**, **Vignesh V.G. Krishnan⋆**, S.A.Foroutan⋆, M.Touhiduzzaman, **Anurag Srivastava**, **Yinghui Wu**, Adam Hahn, Suresh Sindhu. Cyber Physical Security Analytics for Anomalies in Transmission Protection Systems. *The 54th IEEE Industry Applications Society Annual Meeting (IAS)*, 2018 (*Acceptance rate n/a*)

C.35 **Mohammad Hossein Namaki⋆**, **Yinghui Wu**, **Qi Song⋆**, Peng Lin⋆, Tingjian Ge, Discovering Temporal Graph Association Rules. *The 26th ACM International Conference on Information and Knowledge Management (CIKM)*, 2017 (*20.0% acceptance rate*)

C.36 **Mohammad Hossein Namaki⋆**,Peng Lin⋆,**Yinghui Wu**,Event Pattern Discovery by Keywords in Graph Streams.*IEEE International Conference on Big Data*,2017 (*17.8% acceptance rate*)

C.37 **Mohammad Hossein Namaki⋆**, **Keyvan Sasani⋆**, **Yinghui Wu**, Tingjian Ge. BEAMS: Bounded Event Detection over Graph Streams. *International Conference on Data Engineering (ICDE)*, 2017. (*17.7% acceptance rate*)

C.38 **Mohammad Hossein Namaki⋆**,**F A Rezaur Rahman Chuwdhury**,Md Rakibul Islam,Janardhan Rao Doppa,**Yinghui Wu**. Learning to Speed Up Query Planning in Graph Databases. *International Conference on Automated Planning and Scheduling(ICAPS)*, 2017. (*Acceptance rate n/a*)

C.39* **Wenfei Fan**, **Jingbo Xu**, **Yinghui Wu**, Jiaxin Jiang, Zeyu Zheng, Bohan Zhang, Yang Cao, Chao Tian. Parallelizing Sequential Graph Computations. *ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2017.***(Best Paper Award)*** (*19% acceptance rate*)

C.40 **Wenfei Fan**, **Xueli Liu**, **Yinghui Wu**, **Jingbo Xu⋆** and Ping Lu. GRAPE: Parallelizing Sequential Graph Computations. *International Conference on Very Large Data Bases (VLDB)*, 2017. (*Demo track acceptance rate: 30.4%*) ***(Best Demonstration Paper Award)***

C.41 **Qi Song⋆**, **Yinghui Wu**, Xin Luna Dong. Mining Summaries for Knowledge Graph Search. *The IEEE International Conference on Data Mining series (ICDM)*, 2016. (*19.6% acceptance rate; regular research track:* 8.6%) [Journal Version: J.7]

C.42 **Wenfei Fan**, **Yinghui Wu**, Jingbo Xu⋆. Functional Dependencies for Graphs. *ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2016. (*20% acceptance rate*)

C.43 **Shengqi Yang⋆**, Fangqiu Han, **Yinghui Wu**, **Xifeng Yan**. Fast Top-k Search in Knowledge Graphs. *International Conference on Data Engineering (ICDE)*, 2016. (*Acceptance rate n/a; average acceptance rate (last 5 years):* 15.8%)

C.44= Wenfei Fan, **Yinghui Wu**, Jingbo Xu⋆. Adding Counting Quantifiers to Graph Patterns. *ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2016. (*20% acceptance rate*)

C.45= Wenfei Fan,Xin Wang,**Yinghui Wu**,Jingbo Xu⋆. Association Rules with Graph Patterns. *The 41th International Conference on Very Large Data Bases (VLDB)*, 2015. (*21.2% acceptance rate*)

C.46 Predrag T.Tosic, **Yinghui Wu**. Towards Networks of Search Engines and Other Digital

Experts: A Distributed Intelligence Approach. *IEEE International Conference on U-and E-Service Science and Technology*, 2015. (*Acceptance rate n/a*)

C.47 **Bo Zong*****, Yinghui Wu**, Jie Song, Ambuj Singh, Hasan Cam, Jiawei Han, **Xifeng Yan**. Towards Scalable Critical Alert Mining. *The 20th International Conference on Knowledge Discovery and Data Mining (KDD)*, 2014. (*14.6% acceptance rate*)

C.48 **Yinghui Wu**, Shengqi Yang*, Xifeng Yan. Summarizing Answer Graphs Induced by Keyword Queries. *The 40th International Conference on Very Large Data Bases (VLDB)*, 2014. (*20.0% acceptance rate*)

C.49 **Wenfei Fan**, Xin Wang, **Yinghui Wu**, Deng Dong. Distributed Graph Simulation:Impossibility and Possibility. *The 40th International Conference on Very Large Data Bases(VLDB)*, 2014. (*20.0% acceptance rate*)

C.50 **Shengqi Yang*****, Yinghui Wu**, Huan Sun*, **Xifeng Yan**. Schemaless and Structureless Graph Querying. *The 40th International Conference on Very Large Data Bases (VLDB)*, 2014. (*20.0% acceptance rate*)

C.51 **Shengqi Yang*****, Yanan Xie*****, Yinghui Wu**, Tianyu Wu, Huan Sun, Jian Wu, **Xifeng Yan**. SLQ: A User-friendly Graph Querying System, *ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2014 (*25.6% acceptance rate*)

C.52⁼ Wenfei Fan, Xin Wang, **Yinghui Wu**. Querying Big Graphs within Bounded Resources. *ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2014 (*25.6% acceptance rate*)

C.53⁼ Wenfei Fan, Xin Wang, **Yinghui Wu**. Answering Graph Pattern Queries Using Views. *International Conference on Data Engineering (ICDE)*, 2014. (*19.9% acceptance rate*) **(*Best Paper Runner-up*)** [Journal Version: J.12]

C.54⁼ Wenfei Fan, Xin Wang, **Yinghui Wu**. Diversified Top-K Graph Pattern Matching. *The 40th International Conference on Very Large Data Bases (VLDB)*, 2014. (*20.0% acceptance rate*)

C.55 **Yinghui Wu**, Shengqi Yang*, Xifeng Yan. Ontology-based Subgraph Querying.*International Conference on Data Engineering*(ICDE), 2013 (*21.4% acceptance rate*) **(*Best Poster Paper Award*)**

C.56 **Arijit Khan*****, Yinghui Wu**, Charu Aggarwal, Xifeng Yan. NeMa: Fast Graph Search with Label Similarity. *The 39th International Conference on Very Large Data Bases (VLDB)*, 2013. (*22.7% acceptance rate*)

C.57⁼ Wenfei Fan, Xin Wang, **Yinghui Wu**. ExpFinder: Finding Experts by Graph Pattern Matching. *International Conference on Data Engineering (ICDE)*, 2013. (*21.4% acceptance rate*)

C.58⁼ Wenfei Fan, Xin Wang, **Yinghui Wu**. Performance Guarantees for Distributed Reachability Queries. *The 38th International Conference on Very Large Data Bases (VLDB)*, 2012. (*20.3% acceptance rate*)

C.59 **Bo Zong*****, Yinghui Wu**, Ambuj Singh, Xifeng Yan. Inferring the Underlying Structure of Information Cascades. *International Conference on Data Mining* (ICDM), 2012. (*19.9% accep-

tance rate; regular research track: 10.7%)

C.60 **Arijit Khan**[⋆], **Yinghui Wu**, Xifeng Yan. Emerging Graph Queries in Linked Data. *International Conference on Data Engineering (ICDE)*, 2012. (*24.2% acceptance rate*)

C.61[=] Wenfei Fan,Jianzhong Li,Xin Wang,**Yinghui Wu**. Querying Preserving Graph Compression. *ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2012 (*16.6% acceptance rate*)

C.62[=] Wenfei Fan, Jianzhong Li, Jizhou Luo, Zijing Tan, Xin Wang, **Yinghui Wu**. Incremental Graph Pattern Matching. *ACM SIGMOD Conference on Management of Data (SIGMOD)*, 2011. (*23.2% acceptance rate*) [Journal Version: J.13]

C.63[=]Wenfei Fan, Jianzhong Li, Shuai Ma, Nan Tang, **Yinghui Wu**. Adding Regular Expressions to Graph Reachability and Pattern Queries. *International Conference on Data Engineering (ICDE)*, 2011. (*19.8% acceptance rate*) [Journal Version: J.15]

C.64[=] Wenfei Fan, Jianzhong Li, Shuai Ma, Nan Tang, **Yinghui Wu**, Yunpeng Wu. Graph Pattern Matching: From Intractable to Polynomial Time. *The 36th International Conference on Very Large Data Bases (VLDB)*, 2010. (*17.1% acceptance rate*)

C.65[=] Wenfei Fan, Jianzhong Li, Shuai Ma, Hongzhi Wang, **Yinghui Wu**. Graph Homomorphism Revisited for Graph Matching. *The 36th International Conference on Very Large Data Bases (VLDB)*, 2010. (*17.6% acceptance rate*)

C.66[=] Ting Deng, Wenfei Fan, Leonid Libkin, **Yinghui Wu**. On the Aggregation Problem for Synthesized Web Services. *The 13th International Conference on Database Theory (ICDT)*, 2010. (*36.4% acceptance rate*)

C.67 **Wenfei Fan**, **Shuai Ma**, Yanli Hu, Jie Liu, **Yinghui Wu**. Propagating Functional Dependencies with Conditions.*The 34th International Conference on Very Large Data Bases (VLDB)*, 2008. (*17.1% acceptance rate*)

Refereed Workshop Papers

W.1 **Mohammad Hossein Namaki**[⋆], Keyvan Sasani[⋆], **Yinghui Wu**, Assefaw Gebremedhin. Performance Prediction for Graph Queries. *Network Data Analytics at SIGMOD*, 2017.

W.2 **Qi Song**[⋆], Mohammad Hossein Namaki[⋆], Peng Lin[⋆], **Yinghui Wu**. Parallel Graph Summarization for Knowledge Search. *Mining and Learning with Graphs at KDD*, 2017.

Invited Book Chapters

B.1 **Yinghui Wu**, Arijit Khan. Graph Pattern Matching, in Sakr, Sherif, Zomaya, Albert (Eds.), *Encyclopedia of Big Data Technologies*, Springer, 2018

B.2 Davide Mottin, **Yinghui Wu**. Graph Exploration and Search, in Sakr, Sherif, Zomaya, Albert (Eds.), *Encyclopedia of Big Data Technologies*, Springer, 2018

Dissertation

Yinghui Wu. Extending Graph Homomorphism and Simulation for Real-life Graph Matching. Ph.D. thesis, the University of Edinburgh, Edinburgh, UK, 2011.

PATENTS

U.S. Patent Application 407925: "Tracking Provenance in Data Science Scripts". 2020. Avrilia Floratou, Ashvin Agrawal, Mohammad Hossein Namaki, Subramaniam Venkatraman Krishnan, Fotios Psallidas, Yinghui Wu. Provisional filed January 14, 2020.

U.S. Patent Application 62/695,317: "Knowledge graph for real time industrial control system security event monitoring and management". 2019. Jiaxing Pi, Dong Wei, Leandro Pfleger De Aguiar, Yinghui Wu. Provisional filed July 9, 2019.

U.S. Patent Application 62/780,926:"Quality-aware keyword query suggestion and evaluation." 2018.Yinghui Wu, Mohammad Hossein Namaki, Xin Zhang. Provisional filed December 17, 2018.

U.S. Patent US9798774B1: "Graph data search method and apparatus." Wenfei Fan, Xin Wang, Yinghui Wu. Issued October 24, 2017; Provisional filed December 9, 2014.

INVITED TALKS

"Spatiotemporal Graph Neural Networks for PV Data Analytics," the 8th Data Science in Life Science and Engineering Collaboration and Symposium, CWRU-Tohoku University, CWRU, Aug. 2022

"Knowledge-based Event Analysis in Cyber-Physical Data Streams," Data Science in Life Science and Engineering Collaboration and Symposium, CWRU-Tohoko University, Mar. 2021

"Knowledge Graphs, Exploratory Data Systems and Applications," FirstEnergy, Apr. 2021

"Knowledge Graphs, Exploratory Data Systems and Applications," CWRU/NASA AI/DS/ML Workshop, Apr. 2021

"Answering Why-questions for Graph Exploration." *Pacific Northwest National Laboratory*, Richland, WA, Sep. 2020

"Knowledge-based Virtual Assistant for Distributed Grid Control." *American Electric Power*, Webinar, May 2020

"Knowledge and Data Quality." *NSF Convergence-Accelerator All-Group Meeting*, UT Arlington, Arlington, TX, Dec. 2019

"Integrating Physics with Data Analytics for Cyber-Physical Event and Root Cause Analysis." *Workshop on Cyberphysical Security Analytics for the Power Grid*, Princeton University, Princeton, NJ, Oct. 2019

"Towards On-the-Fly Knowledge Bases for Electric Grid Resilience." *NSF Workshop on Real Time Data Analytics for the Power Grid Resiliency*, Portland, OR, Aug. 2018

"From Think Parallel to Think Sequential." Seminar Series in Analysis in Motion, *Pacific North-*

west National Laboratory, Richland, WA, Mar. 2018

"Ensemble-based Bad Data Detection in PMU Data Streams." *NSF Workshop on Data Analytics for the Smart Grid*, Pullman, WA, Aug. 2017

"Answering Pattern Queries Using Views." *Workshop on Network Data Analytics, collocated with SIGMOD Conference.* Chicago, IL, May 2017

"Big Data: State-of-the-art and Open Challenges." *CS Graduate Research Proseminar, Washington State University, Pullman, WA*, Apr. 2017

"Big Graph Analytics: A Journey of Scalability and Usability. " *Pacific Northwest National Laboratory, Richland, WA*, Apr. 2015

"Big Graph Search and Analytics." *Northeast Univ. (Shenyang, Liaoning), Jilin Univ. (Changchun, Jilin), Beihang Univ. (Beijing), China*, Oct. 2014

"Optimizing Service Reachability Against Network Attacks." *Science for Cybersecurity (S4C) workshop*, University of Maryland, College Park, MD, Jun. 2013

"Ontology-based Subgraph Querying." *ICDE Workshop on Graph Data Management*, Brisbane, Australia, Apr. 2013

"Towards Advanced Search in Complex Graphs." *ICDE Workshop on Graph Data Management*, Brisbane, Australia, Apr. 2013

"Ontology-based Subgraph Querying. " *Army Research Laboratory Network Science CTA Technical Meeting*, University of Delaware, Newark, DE, Apr. 2013

"Emerging Graph Queries in Linked Data." *ICDE Tutorial*, Washington DC, Apr. 2012

TEACHING

Instructor (Case Western Reserve University) [7]

CSDS 433: "Database Systems" (online course)   Fall 2021, Summer 2023
An online course developed by the instructor for online MS CS program

CSDS 234: "Structured and Unstructured Data" [8]   Fall 2020 - 2023
21 students (required course for CS and DS major)

CSDS 433: "Database Systems" [9],   Spring 2020 - 2023
30 students, Course Rating: 3.93, Instructor Rating 4.36

DSCI 234: "Structured and Unstructured Data" [10]   Fall 2019
21 students, Course Rating: 4.0, Instructor Rating 4.0 (required course for CS and DS major)

---

[7]CS: Computer Science; MSE: Materials Science & Engineering; EE: Electrical Engineering; DS: Data Science.
[8]Originally DSCI 234
[9]Originally EECS 433
[10]A new course designed by the instructor in fall 2019

Past Courses (Other Institutions)

Instructor

- CPTS 580 "Advanced Databases" (30 students; a new graduate course designed and regularly offered in spring by the instructor), Washington State University, 2015-2018;

- CPTS 415 "Big Data" (40 students; a new undergraduate course designed and regularly offered in fall by the instructor), Washington State University, 2015-2018;

- CPTS 580-2 "Advanced Databases" Online course (designed by instructor and regularly offered in fall to WSU Global Campus), 2017-2019;

- CPTS 415-2 "Big Data" Online course (designed by instructor and regularly offered in fall to WSU Global Campus), 2016-2018;

- CPTS 317 "Formal Language and Automata" (75 students; Undergraduate core course required by CS major), Washington State University, Spring 2018-2019

Co-Instructor

- "Data Mining" (Introduced and co-taught with Xifeng Yan, graduate course; responsibility 70%), University of California Santa Barbara, Fall 2013;

- "Advanced Topics in Web Databases" (regularly offered graduate course in spring, co-taught with Wenfei Fan; responsibility 30%), University of Edinburgh, 2009-2011;

- "Research Topics in Distributed Databases" (graduate course co-taught with Wenfei Fan; responsibility 30%), University of Edinburgh, Spring 2011.

STUDENT
ADVISING

Current Advisees (Case Western Reserve University)

Ph.D. Students

- Hanchao Ma (*passed qualifier*), 2017-present

- Sheng Guan (*passed Thesis Proposal*), 2017-present

- Mengying Wang, 2021-present

- Arafath Nihar, Mingjian Lu, Sameera Nalin Venkat (co-advised with Roger French), 2021-present

- Liangyi Huang (co-advised with Xusheng Xiao), 2021-present

- Ahmad Maroof Karimi (co-advised with Roger French and Mehmet Koyuturk, MSE, *graduated*), 2019-2022

M.S. Students

James Kennelly, CS M.S Thesis	Spring 2021 - 2023
Haolai Che, CS M.S Thesis	Spring 2021 - 2023
Yiyang Bian, CS M.S Thesis	Spring 2021 - 2023
Lei Ruan, CS M.S Thesis	Spring 2020 - 2022
Yunzhou Cao, CS M.S. Thesis	Spring 2020 - 2022
Hao Li, CS M.S. Thesis	Spring 2020 - 2022

Undergraduate Students

Christopher Toomey, CS Major, Research Intern	Summer 2020 - Current
Mocun Ye, CS Major, Research Intern	Summer 2020 - Current

Current Advisees (Other Institutions)

- Arman Ahmad (Ph.D., co-advised with Anurag Srivastava in *PNNL Distinguished Graduate Student Program* [11], EE; responsibility: 40%, *passed qualifier*), WSU, 2017-2022 (Graduated)

Past Students

Graduate Students

- Peng Lin (CS Ph.D. 2021) [*Google*]
- Qi Song, CS Ph.D. 2020, WSU [Applied Research Scientist, *Amazon*]
- Mohammad Hossein Namaki, CS Ph.D. 2020, WSU (graduated with *Best Research Assistant Award* in School of EECS, WSU) [Research Scientist, *Microsoft*]
- Sukhjinder Singh, CS M.S. 2019, WSU [Data Scientist, *Microsoft*]
- Xin Zhang, CS M.S. 2019, WSU [Ph.D. candidate at *University of California, Riverside*]
- Jinling Tao, CS M.S. 2018, WSU [*Mercury Systems (New Jersey)*]
- Mengze Zhou, EE M.S. 2018, WSU [Associate Engineer, *PacifiCorp*]
- Keyvan Sasani, CS M.S. 2017, WSU [Product Engineer, *Bigstream*]
- Giridhar Monoharan, CS M.S. 2017, WSU [Software Developer, *Center for Sustaining Agriculture and Natural Resources, WSU*]
- Mohammed Liaqat, CS M.S. 2011, U of Edinburgh [Senior Engineer, *First Elec. Tech.*]
- Yi Zhang, CS M.S. 2011, U of Edinburgh [Senior Software Engineer, *Servelec Controls*, UK].

Undergraduate Advisees

- Evan McElheny, Marist College, NSF Research Experiences for Undergraduates (REU) program [12] 2019 [president of *Marist College* Computer Society]
- Casey Fleck, Coastal Carolina University, REU 2018 [Software Engineer, *Horry Tele. Coop.*]
- Jialiang Shen, Beijing University of Posts and Telecommunications, Summer Intern 2017 [M.S.,*Carnegie Mellon University*]
- Shayan Monadjemi, Univ. of Texas Dallas, REU 2016 [Ph.D., *Washington Univ. in St. Louis*]
- Rebecca Baumher, University of Pennsylvania, REU 2015 [Software Engineer, *Google*]

Thesis Committees

Ph.D. Committees at CWRU: (1) Sreehari Sankar, (2) Ahmad M. Karimi, (3) Leonard Dervishi, (4) Serhan Yilmaz.

Ph.D. Committees Outside Department: (1) Lei Cai (Texas A&M University, 2020), (2) Xuanming Liu (University of Massachusetts Lowell, 2019), (3) Yuchao Ma (WSU, 2019), (4) Parastoo Alinia (WSU, 2019), (5) Armen Abnousi (WSU, 2018), (6) Priyanka Ghosh (WSU, 2018), (7) Yue Zhang (WSU, 2017), (8) Beiyu Lin (WSU, 2017), (9) Christopher Pereyda (WSU, 2017), (10) Marco Minutoli (WSU, 2017).

---

[11]The student will move to and be supported by PNNL to complete his thesis research from Spring 21.

[12]I served as a faculty mentor for the NSF Smart Environments REU program hosted by the School of EECS, WSU.

M.S. Committees at CWRU: (1) Yantong Li, (2) Tiantong Ji.

M.S. Committees Outside Department: (1) Omkar Kenjale (WSU, 2019), (2) Dipankar Medhi (WSU, 2019), (3) Zayaan Ahmed (WSU, 2019), (4) Arman Ahmad (WSU, 2018), (5) Jinglin Tao (WSU, 2018), (6) Pascal Schmitt (Univ. of Edinburgh, UK, 2016).

Editorial Boards

Associate Editor, ACM Journal of Data and Information Quality (JDIQ)	2018-present

Program Committees

ACM Int'l Conf. on Management of Data (SIGMOD)	2015-2016,2018-2021, 2023
ACM Int'l Conf. on Knowledge Discovery and Data Mining (SIGKDD)	2015-2016,2018-2022
ACM Int'l Conf. on Info. and Knowledge Management(CIKM)	2012-2013,2015,2017,2020,2023
ACM Int'l Conf. on Information Retrieval (SIGIR)	2020
AAAI Conference on Artificial Intelligence (AAAI)	2020, 2022 (**Senior PC**)
Int'l Conf. on Very Large Data Bases (VLDB)	2013-2015,2017,2019-2021, 2023
IEEE Int'l Conf. on Data Engineering (ICDE)	2014,2019-2020, 2023
IEEE Int'l Conf. on Data Mining (ICDM)	2015, 2020
IEEE Int'l Conf. on Big Data (BigData)	2019-2023
Int'l Conf. on Extending Database Technology (EDBT)	2014,2017
Int'l Conf. on Web-Age Information management (WAIM)	2014-2016
Int'l Conf. on Database Systems for Advanced Applications (DASFAA)	2014-2015,2019
SIAM Int'l Conf. on Data Mining (SDM)	2015-2016
European Intelligence and Security Informatics Conference (EISIC)	2013-2015
International Workshop on Graph Data Management and Network Data Analytics	2018-2020
International Workshop on Cloud Data Management	2015
International Workshop on Querying Graph Structured Data	2015-2016

Conference Organization

Session Chair, ACM International Conference on Management of Data (SIGMOD)	2020
Session Chair, International Conference on Very Large Data Bases (VLDB)	2020
Co-organizer, NSF Convergence-Accelerator Workshop (co-located with SIGMOD 2020)	2020
Panelist, NSF Workshop on Smart Grids Big Data	2018
Session Chair, NSF Workshop on Data Analytics for the Power Grid	2018
Session Chair, Workshop on Data Analytics for the Smart Grid	2017

Session Chair, IEEE International Conference on Data Engineering (ICDE)   2013

Journal Referee

ACM Transactions on Database Systems (TODS)

ACM Journal of Data and Information Quality (JDIQ)

Very Large Data Bases (VLDB) Journal[13]

IEEE Transactions on Big Data (TBD)

IEEE Transactions on Knowledge and Data Engineering (TKDE)

IEEE Transactions on Knowledge Discovery from Data (TKDD)

IEEE Transactions on Smart Grid (TSG)

IEEE Transactions on Power Delivery (PWRD)

IEEE Transactions on Industrial Informatics (TII)

IEEE Transactions on Parallel and Distributed Systems (TPDS)

PLOS One

Distributed and Parallel Databases (DPD)

Information Systems (IS)

World Wide Web (WWW) Journal

Grant Review

Panel Member, National Science Foundation	2014,2017,2019-present
Panel Member, Department of Homeland Security Centers of Excellence	2017
International Panel Member, Hong Kong Research Grants Council	2018,2020

INSTITUTIONAL
SERVICE

Services to Case Western Reserve University

Department Service

- *Program Representative* (Data Science); Undergraduate Affairs Committee   2020-Present
  - Coordinated the preparation of ABET Pilot Accreditation of Data Science & Analytics Undergraduate Degree Program at CDS [14];
  - Regularly provide information to navigators, incoming first-year students, senior students, and be responsible for other related information requests;
  - Promoted CWRU DS program in connection to ABET; advised in ABET DS working group on sculpting new ABET accreditation criteria for DS programs in U.S. universities;
  - Outreaching activities: hosted visiting students, scholars and industry partners; promoted DS programs and faculty to industry partners and stakeholders (in collaboration with Great Lakes Energy Institute).

---

[13]This bi-monthly journal of VLDB Endowment has one issue devoted to selected papers from VLDB Conference.
[14]The Data Science Program at CDS received positive feedback from ABET visit in October 2021, and is expected to be the first ABET accredited Data Science program in US

- *Member*, Graduate Education Committee 2019-2020
  - Participated in regular discussions and advised on educational issues and strategic initiatives, e.g., curriculum design and promotion, admission criteria, and evaluation policies.
  - Facilitated the coordination and mapping of Accreditation Board for Engineering and Technology (ABET) criteria to existing CS and DS curriculum.
- *Faculty Advisor* of Data Science Major 2019-Present
  - Provided guidance and suggestions to students on course selection, time management, resume and career development (e.g., internships), among others.
- *Faculty Advisor* of BS in CS Major (Database and Data Mining track) 2019-Present
  - Advise on similar topics as above for CS students; advise on specific research-related topics in database and data mining.

School Service

- *Member*, Undergraduate Education Committee, Case School of Engineering [15] 2020 - Present
  - Participated in regular committee meetings; responsible for facilitating and advising on various undergraduate education issues, including course and program actions, curriculum adjustment, evaluation policies and student/faculty feedback solicitation.

Previous Services (Other Institutions)

School Service

- **Member**, Graduate Studies Committee, School of EECS, WSU 2017-2019

University Service

- **Faculty Mentor**, Louis Stokes Alliances for Minority Participation; WSU 2017-2019

---

[15]This appointment started from Fall 2020.