

Yinghui Wu

CONTACT INFORMATION

Department of Computer and Data Sciences
Case Western Reserve University
Olin 515, 2101 Martin Luther King Jr. Dr
Cleveland, OH 44106

Office: (216) 368-8829
Phone: (805) 679-1982
E-mail: yxw1650@case.edu
Homepage: yinghuwu.github.io

RESEARCH INTERESTS

Data and Knowledge Management, Data Systems, Information Network, Information quality

EDUCATION

The University of Edinburgh, Edinburgh, UK **2011**

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

Peking University, Beijing, China **2007**

B.S. in Computer Science

Peking University, Beijing, China **2007**

Bachelor of Economics (Double Major)

ACADEMIC EXPERIENCE

Case Western Reserve University, Cleveland, OH **2019 - present**

Assistant Professor, Case School of Engineering

Pacific Northwest National Laboratory, Richland, WA USA **2017 - present**

Staff Scientist (Joint), Advanced Computing, Mathematics, and Data Division

Washington State University, Pullman, WA **2014 - 2019**

Assistant Professor, School of Electrical Engineering and Computer Science

University of California Santa Barbara, Santa Barbara, CA USA **2011 - 2014**

Research Scientist, Department of Computer Science

PUBLICATIONS

**: authors are listed in alphabetical order (equal contribution). ★: student*

Conference Papers

C.1 Peng Lin^{*}, Qi Song^{*}, Yinghui Wu, Jiaying Pi. Repairing Entities using Star Constraints in Multirelational Graphs. Int.Conf.on Data Engineering (ICDE), 2020.

C.2 Hanchao Ma^{*}, Morteza Alipour Langouri, Yinghui Wu, Fei Chiang, Jiaying Pi. Ontology-based Entity Matching in Attributed Graphs. The 45rd International Conference on Very Large Data Bases (VLDB), 2019

C.3 Sheng Guan^{*}, Hanchao Ma^{*}, Yinghui Wu, Jiaying Pi. Attribute-Driven Backbone Discovery. The 25th Int. Conf. on Knowledge Discovery and Data Mining (KDD), 2019

C.4 Mohammad Hossein Namaki^{*}, Qi Song^{*}, Yinghui Wu. NAVIGATE: Explainable Visual Graph Exploration by Examples. ACM SIGMOD Conference on Management of Data (SIGMOD), 2019

C.5 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

C.6 Qi Song^{*}, Mohammad Hossein Namaki^{*}, **Yinghui Wu**. Answering Why-Questions for Sub-graph Queries in Multi-Attributed Graphs. Int. Conf. on Data Engineering (ICDE), 2019.

- C.7 Xuanming Liu, Tingjian Ge, **Yinghui Wu**. Finding Densest Lasting Subgraphs in Dynamic Graphs: a Stochastic Approach Int. Conf. on Data Engineering (ICDE), 2019.
- C.8 Mohammad Hossein Namaki*, **Yinghui Wu**, Xin Zhang*. GExp: Cost-aware Graph Exploration with Keywords. ACM SIGMOD Conference on Management of Data (SIGMOD), 2018.
- C.9 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
- C.10 Peng Lin*, Qi Song*, Jialiang Shen, **Yinghui Wu**. Discovering Graph Patterns for Fact Checking in Knowledge Graphs 23rd International Conf. on Database Systems for Advanced Applications (DASFAA), 2018
- C.11 Keyvan Sasani*, Mohammad Hossein Namaki*, **Yinghui Wu**, Assefaw Gebremedhin. Multi-metric Graph Query Performance Prediction 23rd International Conf. on Database Systems for Advanced Applications (DASFAA), 2018
- C.12 Mohammad Hossein Namaki*, **Yinghui Wu**, Xin Zhang*. Diversified Keyword Expansion on Multi-labeled Graphs. APWeb-WAIM Joint Conference on Web and Big Data, 2018
- C.13 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
- C.14 Wenfei Fan, Yang Cao, Jingbo Xu, Wenyuan Yu, **Yinghui Wu**, Chao Tian, Jiabin Jiang, Bohan Zhang. From Think Parallel to Think Sequential. SIGMOD Record, 2018
- C.15 Discovering Temporal Graph Association Rules The 26th ACM International Conference on Information and Knowledge Management (CIKM), 2017 Mohammad Hossein Namaki*, **Yinghui Wu**, Qi Song*, Peng Lin*, Tingjian Ge
- C.16 Mohammad Hossein Namaki*, Peng Lin*, **Yinghui Wu** Event Pattern Discovery by Keywords in Graph Streams IEEE International Conference on Big Data (IEEE BigData), 2017
- C.17 Mohammad Hossein Namaki*, Keyvan Sasani*, **Yinghui Wu**, Tingjian Ge. BEAMS: Bounded Event Detection over Graph Streams. Int.Conf.on Data Engineering (ICDE), 2017.
- C.18 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*, 2017.
- C.19 Wenfei Fan, Jingbo Xu, **Yinghui Wu**, Jiabin Jiang, Zeyu Zheng, Bohan Zhang, Yang Cao, Chao Tian, Parallelizing Sequential Graph Computations. ACM SIGMOD Conference on Management of Data (SIGMOD), 2017
- C.20 Wenfei Fan, Xueli Liu, **Yinghui Wu**, Jingbo Xu and Ping Lu, GRAPE: Parallelizing Sequential Graph Computations. Int. Conf. on Very Large Databases (VLDB), 2017.
- C.21 Qi Song*, **Yinghui Wu**, Xin Luna Dong, Mining Summaries for Knowledge Graph Search. The IEEE International Conference on Data Mining series (ICDM), 2016.
- C.22 Wenfei Fan, **Yinghui Wu***, Jingbo Xu, Functional Dependencies for Graphs. ACM SIGMOD Conference on Management of Data (SIGMOD), 2016.
- C.23 Shengqi Yang, Fangqiu Han, **Yinghui Wu**, Xifeng Yan, Fast Top-k Search in Knowledge Graphs. Int. Conf. on Data Engineering (ICDE), 2016.
- C.24 Wenfei Fan, **Yinghui Wu***, Jingbo Xu, Adding Counting Quantifiers to Graph Patterns. ACM SIGMOD Conference on Management of Data (SIGMOD), 2016.

- C.25 Wenfei Fan, Xin Wang, **Yinghui Wu***, Jingbo Xu. Association Rules with Graph Patterns. The 41th Int. Conf. on Very Large Databases (VLDB), 2015.
- C.26 Predrag T.Tosic, **Yinghui Wu**. Towards networks of search engines and other digital experts: a distributed intelligence approach. IEEE Int.Conf.on u-and e-Service Science and Technology, 2015.
- C.27 **Yinghui Wu**, Shengqi Yang, Xifeng Yan. Summarizing Answer Graphs Induced by Keyword Queries. Proc. of the 40th Int. Conf. on Very Large Databases (VLDB), 2014.
- C.28 Bo Zong, **Yinghui Wu**, Jie Song, Ambuj Singh, Hasan Cam, Jiawei Han, Xifeng Yan. Towards Scalable Critical Alert Mining. The 20th Int. Conf. on Knowledge Discovery and Data Mining (KDD), 2014.
- C.29 Wenfei Fan, Xin Wang, **Yinghui Wu**, Deng Dong. Distributed Graph Simulation: Impossibility and Possibility. Proc. of the 40th Int. Conf. on Very Large Databases (VLDB), 2014.
- C.30 Wenfei Fan, Xin Wang, **Yinghui Wu***. Querying Big Graphs within Bounded Resources. ACM SIGMOD Conference on Management of Data (SIGMOD), 2014
- C.31 Shengqi Yang, **Yinghui Wu**, Huan Sun, Xifeng Yan. Schemaless and Structureless Graph Querying. Proc. of the 40th Int. Conf. on Very Large Databases (VLDB), 2014.
- C.32 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
- C.33 Wenfei Fan, Xin Wang, **Yinghui Wu***. Answering Graph Pattern Queries Using Views. Int. Conf. on Data Engineering (ICDE), 2014.
- C.34 Wenfei Fan, Xin Wang, **Yinghui Wu***. Diversified Top-K Graph Pattern Matching. Proc. of the 40th Int. Conf. on Very Large Databases (VLDB), 2014.
- C.35 **Yinghui Wu**, Shengqi Yang, Xifeng Yan. Ontology-based Subgraph Querying. Int. Conf. on Data Engineering (ICDE), 2013.
- C.36 Arijit Khan, **Yinghui Wu**, Charu Aggarwal, Xifeng Yan. NeMa: Fast Graph Search with Label Similarity. Proc. of the 39th Int. Conf. on Very Large Databases (VLDB), 2013.
- C.37 Wenfei Fan, Xin Wang, **Yinghui Wu***. ExpFinder: Finding Experts by Graph Pattern Matching. Int. Conf. on Data Engineering (ICDE), 2013.
- C.38 Wenfei Fan, Xin Wang, **Yinghui Wu***. Performance Guarantees for Distributed Reachability Queries. Proc. of the 38th Int. Conf. on Very Large Databases (VLDB), 2012.
- C.39 Bo Zong, **Yinghui Wu**, Ambuj Singh, Xifeng Yan. Inferring the Underlying Structure of Information Cascades. Int. Conf. on Data Mining (ICDM), 2012.
- C.40 Arijit Khan, **Yinghui Wu**, Xifeng Yan. Emerging Graph Queries in Linked Data. Int. Conf. on Data Engineering (ICDE), 2012.
- C.41 Wenfei Fan, Jianzhong Li, Xin Wang, **Yinghui Wu***. Querying Preserving Graph Compression. ACM SIGMOD Conference on Management of Data (SIGMOD), 2012
- C.42 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.
- C.43 Wenfei Fan, Jianzhong Li, Shuai Ma, Nan Tang, **Yinghui Wu***. Adding Regular Expressions to Graph Reachability and Pattern Queries. Int. Conf. on Data Engineering (ICDE), 2011.
- C.44 Wenfei Fan, Jianzhong Li, Shuai Ma, Nan Tang, **Yinghui Wu***, Yunpeng Wu. Graph

Pattern Matching: From Intractable to Polynomial Time. Proc. of the 36th Int. Conf. on Very Large Databases (VLDB), 2010.

C.45 Wenfei Fan, Jianzhong Li, Shuai Ma, Hongzhi Wang, **Yinghui Wu***. Graph Homomorphism Revisited for Graph Matching. Very Large Data Bases (Journal Track), 2010.

C.46 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.

C.47 Wenfei Fan, Shuai Ma, Yanli Hu, Jie Liu, **Yinghui Wu***. Propagating Functional Dependencies with Conditions. Proc. of the 34th Int. Conf. on Very Large Databases (VLDB), 2008.

Journal Papers

J.1 Peng Lin*, Qi Song*, **Yinghui Wu**, Jiaying Pi. Discovering Patterns for Fact Checking in Knowledge Graphs. ACM Journal of Data and Information Quality (JDIQ), 2019

J.2 Peng Lin*, Qi Song*, **Yinghui Wu**. Fact Checking in Knowledge Graphs with Ontological Subgraph Patterns Data Science & Engineering (DSE) (invited), 2019

J.3 Qi Song*, **Yinghui Wu**, Peng Lin*, Xin (Luna) Dong, Hui Sun. Mining Summaries for Knowledge Graph Search IEEE Transactions on Knowledge and Data Engineering (TKDE). 2018

J.4 Mengze Zhou*, Yuhui Wang, Anurag K Srivastava, **Yinghui Wu**, Paramarshi Banerjee. Ensemble Based Algorithm For Syncrophasor Data Anomaly Detection. IEEE Transactions on Smart Grid, 2018

J.5 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.6 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.7 Wenfei Fan, Xin Wang, **Yinghui Wu***. Answering graph query using views. IEEE Transactions on Knowledge and Data Engineering (TKDE), 2015.

J.8 Wenfei Fan, Xin Wang, **Yinghui Wu***. Incremental Graph Pattern Matching. ACM Transactions on Database Systems (TODS), 2013.

J.9 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, 2012

J.10 Ting Deng, Wenfei Fan, Leonid Libkin, **Yinghui Wu***. On the Aggregation problem for Synthesized Web Services. Journal of Computer and System Science (JCSS). 2011

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* and **Yinghui Wu** Parallel Graph Summarization for Knowledge Search. Workshop on Mining and Learning with Graphs at KDD 2017

Book Chapters

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

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

AWARDS AND
HONORS

2019 VLDB Distinguish Reviewer
2018 ACM SIGMOD Research Highlight Award [C.9]
2017 ACM SIGMOD Best Paper Award [C.14]
2017 VLDB Best Demo Award, [C.15]
2015 Google Faculty Research Award
2014 ICDE Best Paper Runner Up, 2014 [C.28]
2013 ICDE Best Poster Award, 2013 [C.30]

SERVICE

Editorial Board

ACM Journal of Data and Information Quality

Program Committee

ACM SIGMOD International Conference on Management of Data (SIGMOD) 2015-2016,2018
ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD) 2015-2016,2018
International Conference on Very Large Data Bases (VLDB) 2013-2015,2017,2019
IEEE International Conference on Data Engineering (ICDE) 2014,2019-2020
ACM Int'l Conf. on Information and Knowledge Management(CIKM) 2012-2013,2015,2017
IEEE International Conference on Data Mining (ICDM) 2015
International Conference on Extending Database Technology (EDBT) 2014,2017
International Conference on Web-Age Information management (WAIM) 2014-2016
International Conference on Database Systems for Advanced Applications (DASFAA) 2014-2015
SIAM International Conference on Data Mining (SDM) 2015-2016
International Workshop on Cloud Data Management (CloudDB) 2014
International Workshop on Querying Graph Structured Data (GraphQ) 2015-2016

Journal Referee

Journal of Computer and System Sciences, ACM Transactions on Database Systems, Distributed and Parallel Databases, IEEE Transactions on Parallel and Distributed Systems, IEEE Transactions on Knowledge and Data Engineering, Information System, VLDB Journal, WWW Journal

Panelist

NSF Grant Review Panel, Hong Kong Research Grants Council

INVITED TALKS

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. Pacific Northwest National Laboratory, Richland, WA, Mar. 2018

Answering Pattern Queries Using Views. SIGMOD workshop on Network Data Analytics (NDA), May. 2019

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

Big Graph Search and Analytics. Northeast Univ., Jilin Univ., Beihang Univ., 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. The 29th ICDE conference, GDM workshop Brisbane, Australia, Apr. 2013

Towards advanced search in complex graphs. The 29th ICDE conference, GDM workshop, Brisbane, Australia, Apr. 2013

Ontology-based subgraph querying. ARL Network Science CTA Technical Meeting, University of Delaware, Newark, DE, Apr. 2013

Emerging Graph Queries in Linked Data. IEEE International Conference on Data Engineering, Washington DC, Apr. 2012

TEACHING

Instructor

Advanced Databases (Graduate), CWRU, Spring 2020

Structured and Unstructured Data (Undergraduate), CWRU, Fall 2019

Advanced Databases (Graduate), Washington State Univ., Spring 2015-2018

Big Data (Undergraduate), Washington State Univ., Fall 2015-2018

Formal Language and Automata (Undergraduate), Washington State Univ., Spring 2018-2019

Co-organizer & Co-Teaching

Graduate Seminars on Data Mining, UC Santa Barbara, Fall 2013

Advanced Topics in Web Databases, Univ. of Edinburgh, Spring 2009, Spring 2010, Spring 2011

Topics in Distributed Databases, Univ. of Edinburgh, Spring 2009, Spring 2010, Spring 2011

CURRENT AND

FORMER STUDENTS

Ph.D.

Qi Song, Mohammad Hossein Namaki, Peng Lin, Hanchao Ma, Sheng Guan, Arman Ahmed

Masters

Xin Zhang, Sukhjinder Singh, Keyvan Sasani, Giridhar Monoharan, Mengze Zhou, Mohammed Saad Liaqat (Univ. of Edinburgh), Yi Zhang (Univ. of Edinburgh).

Undergraduate

Evan McElheny (NSF REU), Casey Fleck (NSF REU), Shayan Monadjemi (NSF REU), Rebecca Baumher (NSF REU), Jialiang Shen.

Thesis/Ph.D. Committee

Lei Cai, Armen Abnoui, Yuchao Ma, Priyanka Ghosh, Pascal Schmitt (Univ. of Edinburgh), Xuanming Liu (Univ. of Massachusetts Lowell).