

HUIPING CAO, Ph.D.

Associate Professor
Department of Computer Science
College of Arts and Sciences
New Mexico State University (NMSU)
Las Cruces, NM 88003, USA

Phone: (575)646-4600
Email: hcao@cs.nmsu.edu
Home Page: <http://www.cs.nmsu.edu/~hcao>

1 RESEARCH INTERESTS

- Data mining: time series (e.g., spatiotemporal data, sensor data) mining, graph (e.g., social networks) mining
- Applied machine learning: probabilistic graphical models, deep neural networks
- Data engineering: graph search, data integration
- Application domains: smart grids, high performance computing, agriculture

2 EDUCATION

- **Doctor of Philosophy** in Computer Science (09/2002 - 11/2007)
The University of Hong Kong, Hong Kong
Thesis title: Pattern Discovery from Spatiotemporal Data
Supervisors: Dr. David W. Cheung, Dr. Nikos Mamoulis
- **Master of Science** in Computer Science (09/1999 - 07/2002)
School of Information, Renmin University of China, Beijing, China
- **Bachelor of Science** in Management Information Systems (09/1995 - 07/1999)
School of Information, Renmin University of China, Beijing, China

3 APPOINTMENTS

- **Associate Professor** (08/2016 - present)
Department of Computer Science, New Mexico State University
Las Cruces, New Mexico
- **Assistant Professor** (08/2010 - 08/2016)
Department of Computer Science, New Mexico State University
Las Cruces, New Mexico
- **Postdoctoral Research Fellow** (11/2009 - 07/2010)
National Center for Ecological Analysis and Synthesis,
University of California Santa Barbara
Santa Barbara, California
- **Postdoctoral Research Associate** (06/2007 - 11/2009)
Computer Science & Engineering, Arizona State University
Tempe, Arizona

4 TEACHING EXPERIENCE

4.1 New Mexico State University

Department of Computer Science

Las Cruces, New Mexico

- C S 272/463: Introduction to Data Structures (undergraduate required course):
Spring 2012, Spring 2013, Fall 2013, Spring 2014, Fall 2014, Spring 2015, Fall 2015, Spring 2019, Fall 2019, Spring 2020
- CS 343/493: Algorithm Design and Implementation:
Fall 2018
- C S 371/468: Software Development
Fall 2017
- C S 479/579: Special Topic on Data Mining (undergraduate and graduate cross-listed course):
Fall 2011, Fall 2012
- CS 479: Special Topic on Algorithm Design and Implementation
Spring 2016
- C S 487/519: Applied Machine Learning I (undergraduate and graduate cross-listed course):
Fall 2018, Spring 2020
- C S 488/508: Introduction to Data Mining (undergraduate and graduate cross-listed course):
Fall 2014, Fall 2015, Fall 2016
- C S 482/502: Database Management Systems I (undergraduate and graduate cross-listed course):
Fall 2010, Fall 2011, Fall 2012, Fall 2013, Fall 2016
- C S 582: Database Management Systems II (graduate course):
Spring 2011, Spring 2012, Spring 2013, Spring 2014, Spring 2015, Spring 2016

4.2 Arizona State University

Computer Science & Engineering,

School of Computing, Informatics, and Decision Systems Engineering

Ira Fulton School of Engineering, Arizona State University

Tempe, Arizona

- CPI 310: Information and Data Management (certificate course):
Fall 2008

5 ADVISING AND MENTORING

- Postdoctoral researcher mentoring
 - 1). Milan Biswal (Postdoctoral research associate, 08/2014 - 08/2016)
- Doctoral committee chair (5 in total, in reverse order of the starting time)
 - 1). Huiying Chen (08/2020 - present)
 - 2). Erick Draayer (01/2019 - present)
 - 3). Edgar Ceh Varela (01/2017 - present)
 - 4). Qixu Gong (08/2015 - present)
 - 5). Yifan Hao (08/2014 - present)
 - 6). Chuan Hu (08/2012 - 05/2017, graduated in May 2017)
Thesis title: *Discovering Influence Relationships from Graph Structured Data*
- Master's committee chair (11 in total, in reverse order of the graduation time)
 - 1). Beepana Pokharel (MSc, 01/2020-present)

- 2). Ramin Zahedi Darshoori (graduated in 06/2020)
Project title: *Neural Network Based Approaches to Mobile Target Localization and Tracking Using Unmanned Aerial*
 - 3). Ashley Michalenko (graduated in 12/2018)
Project title: *A System to Streamline Data Analysis on Engineering Coursework*
 - 4). Jacob Voldez (graduated in 12/2018)
Project title: *Analysis of the Effects of Blind Review Policy in Computer Science Conferences*
 - 5). Panika Valecha (graduated in 07/2018)
Project title: *Analysis of Storage Error Events for High Performance Computing Systems*
 - 6). Raul Alvarado Garcia (graduated in 12/2016)
Project title: *Finding Shortest Path using Landmark Index on Neo4*
 - 7). Mitra Solgi (graduated in 12/2015)
Project title: *Classification of Healthcare Data*
 - 8). Mahmoud Sharifi (graduated in 05/2014)
Project title: *Effect of Feature Extraction Techniques on Time Series Clustering and a Proposed Idea for Semi-supervised Clustering*
 - 9). Abdulwart Heshek (graduated in 12/2013)
Project title: *Computing Alliance for Hispanic-Serving Institutions (CAHSI) Ontology Modeling*
 - 10). Yifan Hao (graduated in 05/2013)
Thesis title: *Keyword Search over Graphs using MapReduce*
 - 11). Aditya Madadi (graduated in 12/2012)
Project title: *Approaches to Extract Relationships among Attribute Values from Spatiotemporal Data*
- Undergraduates directly mentored in research projects
(15 in total, in reverse order of the starting time)
Yuxi Wang (05/2020 - present), Wen Yang (04/2020 - present), Jiefei Liu (08/2018 - 05/2020), Omar Navarro (08/2019 - 01/2020), Caitlin Ard (08/2018 - 12/2018), Michael Meerbott (12/2017 - 05/2018), Ian Goetting (12/2017 - 05/2018), Brett Pelkey (06/2015 - 05/2017), Erick Draayer (06/2015 - 12/2015), Harold Hughs (06/2015 - 10/2015), Stefan Ceballos (01/2015 - 01/2016), Erik Ness (01/2015 - 05/2015), Chaomin Ke (01/2012 - 05/2012), Zhe Xie (01/2012 - 05/2012), Yangpai Liu (01/2011-05/2011),

6 PUBLICATIONS

** - undergraduate students whom I advise in research projects;

* - graduate students or postdoctoral researchers whom I advise or work closely with.

6.1 Journal Articles

- J1. Yifan Hao*, **Huiping Cao**, Abdullah Mueen, Sukumar Brahma: Identify Significant Phenomenon-specific Variables for Multivariate Time Series. Accepted to be published in IEEE Trans. Knowl. Data Eng. (TKDE). <https://doi.org/10.1109/TKDE.2019.2934464>
- J2. Sukumar Brahma, Rajesh Kavasseri, **Huiping Cao**, Nilanjan Ray Chaudhuri, Theodoros Alexopoulos, and Yinan Cui: Real Time Identification of Dynamic Events in Power Systems using PMU data, and Potential Applications - Models, Promises, and Challenges. In IEEE Transactions on Power Delivery, Volume: 32, Issue: 1, 2017. Pages 294 - 301. <http://dx.doi.org/10.1109/TPWRD.2016.2590961>.
- J3. Francisco Natividad*, Russell Y. Folk, William Yeoh, and **Huiping Cao**: On the Use of Off-the-Shelf Machine Learning Techniques to Predict Energy Demands of Power TAC Consumers. In Springer Agent Mediated Electronic Commerce, pages 112-126, Feb. 2017. https://link.springer.com/chapter/10.1007%2F978-3-319-54229-4_8

- J4. Chuan Hu* and **Huiping Cao**: Aspect-Level Influence Discovery from Graphs. *IEEE Trans. Knowl. Data Eng. (TKDE)* 28(7): 1635-1649 (2016) <http://dx.doi.org/10.1109/TKDE.2016.2538223>.
- J5. Phani Harsha Gadde, Milan Biswal*, Sukumar Brahma, and **Huiping Cao**, Efficient Compression of PMU Data in WAMS, *IEEE Transactions on Smart Grid.* 7(5): 2406 - 2413, 2016. <http://dx.doi.org/10.1109/TSG.2016.2536718>.
- J6. Milan Biswal*, Sukumar M. Brahma, and **Huiping Cao**. Supervisory Protection and Automated Event Diagnosis using PMU Data, *IEEE Transactions on Power Delivery (TPD)*, 31 (4), pages 1855-1863 (2016). <http://dx.doi.org/10.1109/TPWRD.2016.2520958>.
- J7. Mohammed N. Sawalhah, Andres F. Cibils, Aditya Madadi*, **Huiping Cao**, Dawn M. Vanleeuwen, Jerry L. Holechek, Christina M. Black Rubio, Robert L. Wesley, Rachel L. Endecott, Travis J. Mulliniks, and Mark K. Petersen: Forage and weather influence day- vs. nighttime cow behavior and calf weaning weights on rangeland Rangeland Ecology & Management. In *Rangeland Ecology & Management* 69 (2016), 134 - 143.
- J8. James Obert*, Inna Pivkina, Hong Huang, **Huiping Cao**: Proactively applied encryption in multipath networks. In *Computers & Security*, Volume 58, 106-124, May 2016. <http://www.sciencedirect.com/science/article/pii/S0167404815001960>.
- J9. Mohammed N. Sawalhah, Andres F. Cibils, Chuan Hu*, **Huiping Cao**, and Jerry L. Holechek: Animal-Driven Rotational Grazing Patterns on Seasonally Grazed New Mexico Rangeland. In *Rangeland Ecology & Management* 67:710-714 (November 2014). <http://dx.doi.org/10.2111/REM-D-14-00047.1>
- J10. Yifan Hao*, **Huiping Cao**, Chuan Hu*, Kabi Bhattarai*, Satyajayant Misra: K-anonymity for social networks containing rich structural and textual information. *Social Netw. Analys. Mining* 4(1): 222-261, August (2014). <http://link.springer.com/article/10.1007%2Fs13278-014-0223-3>
- J11. Jingyu Han, Kejia Chen, Zhiming Ding, **Huiping Cao**: An efficient location reporting and indexing framework for urban road moving objects. *Distributed and Parallel Databases* 32(2): 271-311 (2014), Nov. 2013. <http://link.springer.com/article/10.1007%2Fs10619-013-7135-5>
- J12. Om P. Dahal*, Sukumar M. Brahma, **Huiping Cao**: Comprehensive Clustering of Disturbance Events Recorded by Phasor Measurement Units. In *IEEE Transactions on Power Delivery*, Volume: 29 (3): 1390 - 1397, June 2014. <http://dx.doi.org/10.1109/TPWRD.2013.2285097>
- J13. James Obert*, **Huiping Cao**, Hong Huang: Determination of Multipath Security Using Efficient Pattern Matching. In *International Journal of Computer Science and Information Security (IJCSIS)*, 11(11), 24-33 (2013). <http://www.scribd.com/doc/190070743/Determination-of-Multipath-Security-Using-Efficient-Pattern-Matching>
- J14. **Huiping Cao**, Shawn Bowers, Mark P. Schildhauer: Database Support for Enabling Data-Discovery Queries over Semantically-Annotated Observational Data. In *LNCS Transactions on Large-Scale Data- and Knowledge-Centered Systems (TLDKS)*, 6: 198-228 (2012). http://link.springer.com/chapter/10.1007%2F978-3-642-34179-3_7
- J15. K. Selçuk Candan, **Huiping Cao**, Yan Qi, and Maria Luisa Sapino: System Support for Exploration and Expert Feedback in Resolving Conflicts during Integration of Metadata. In *The VLDB Journal*, 17(6): 1407-1444, 2008. <http://www.springerlink.com/content/e63571635m577344/>
- J16. **Huiping Cao**, Nikos Mamoulis, and David W. Cheung: Discovery of Periodic Patterns in Spatiotemporal Sequences. In *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 19(4): 453-467, 2007. <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=4118704>
- J17. **Huiping Cao**, Shan Wang, and Lingwei Li: Location Dependent Query in a Mobile Environment. In *Journal of Information Sciences*, 154(1-2): 71-83, 2003. <http://www.sciencedirect.com/science/article/pii/S0020025503000355>

6.2 Refereed Conference Papers

- C1. Yifan Hao* and **Huiping Cao**: A new attention mechanism to classify multivariate time series. Accepted to be published in *Proc. Intl. Joint Conf. on Artificial Intelligence (IJCAI)* 2020.

- C2. Qixu Gong*, Jiefei Liu** and **Huiping Cao**: CSQ System: A System to Support Constrained Skyline Queries on Transportation Networks (Demo paper). Accepted to be published in Proc. of IEEE Intl. Conf. on Data Engineering (ICDE), 2020.
- C3. Ramin Zahedi*, Edgar Ceh-Varela*, Robert Selje II, Liang Sun, and **Huiping Cao**: Neural Network Based Approaches to Mobile Target Localization and Tracking Using Unmanned Aerial Vehicles. In Proc. of American Institute of Aeronautics and Astronautics (AIAA) SciTech Conference, Orlando, FL, Jan 10-15, 2020. <https://doi.org/10.2514/6.2020-0392>
- C4. Edgar Ceh-Varela* and **Huiping Cao**: Recommending Packages of Multi-criteria Items to Groups. In Proc. of IEEE Intl. Conf. on Web Services (ICWS) 2019: 273-282. (Acceptance rate: 18%) <https://doi.org/10.1109/ICWS.2019.00054>
- C5. Qixu Gong*, **Huiping Cao**, Parth Nagarkar: Skyline Queries Constrained by Multi-cost Transportation Networks. In Proc. of IEEE Intl. Conf. on Data Engineering (ICDE), 2019: 926-937. <https://doi.org/10.1109/ICDE.2019.00087>
- C6. Chuan Hu*, **Huiping Cao**, Qixu Gong*: Sub-Gibbs Sampling: a New Strategy for Inferring LDA. In Proc. of IEEE Intl. Conf. on Data Mining (ICDM 2017), 907-912. <https://doi.org/10.1109/ICDM.2017.113>. (Overall acceptance rate: 19.9%).
- C7. Milan Biswal*, Yifan Hao*, Phillip Chen, Sukumar Brahma, **Huiping Cao**, and Phillip DeLeon, Signal Features for Classification of Power System Disturbances using PMU Data. In Proc. of 19th Power Systems Computation Conference (PSCC), 20-24 June 2016, Genoa, Italy. (Acceptance rate: N.A.) <http://dx.doi.org/10.1109/PSCC.2016.7540867>.
- C8. Michael Brown, Milan Biswal*, Sukumar Brahma, Satish J Ranade, and **Huiping Cao**, Characterizing and Quantifying Noise in PMU data, Proc. IEEE PES General Meeting, 17-21 July 2016, Boston. (Acceptance rate: N.A.) <https://doi.org/10.1109/PESGM.2016.7741972>.
- C9. Chuan Hu* and **Huiping Cao**: Discovering Time-evolving Influence from Dynamic Heterogeneous Graphs. In Proc. of IEEE International Conference on Big Data 2015, 2253-2262. (Acceptance rate: 42.8%). <http://dx.doi.org/10.1109/BigData.2015.7364014>
- C10. Yifan Hao*, **Huiping Cao**, Yan Qi, Chuan Hu*, Sukumar Brahma, and Jingyu Han: Efficient Keyword Search on Graphs using MapReduce. In Proc. of IEEE International Conference on Big Data 2015, 2871-2873. <http://dx.doi.org/10.1109/BigData.2015.7364106>
- C11. Om P. Dahal, **Huiping Cao**, Sukumar M. Brahma: Evaluating Performance of Classifiers for Supervisory Protection using Disturbance Data from Phasor Measurement Units. In IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2014: 1-6. Oct. 12-15, 2014. (Acceptance rate: N.A.) <http://dx.doi.org/10.1109/ISGTEurope.2014.7028892>.
- C12. James Obert*, Inna Pivkina, Hong Huang, **Huiping Cao**: Dynamically Differentiated Multipath Security in Fixed Bandwidth Networks. In Military Communications Conference (MILCOM 2014), 88-93. Oct. 6 - 8, 2014. (Acceptance rate: N.A.) <http://dx.doi.org/10.1109/MILCOM.2014.22>
- C13. Chuan Hu*, **Huiping Cao**, Chaomin Ke**: Detecting Influence Relationships from Graphs. Proc. of SIAM Data Mining, SDM 2014:821-829. (Acceptance rate: N.A.) <http://dx.doi.org/10.1137/1.9781611973440.94>
- C14. Yangpai Liu**, **Huiping Cao**, Yifan Hao*, Peng Han, Xinda Zeng**: Discovering Context-aware Influential Objects. In Proc. of SIAM Data Mining, SDM 2012:780-791. (Acceptance rate: 27%) <http://siam.omnibooksonline.com/2012datamining/data/papers/237.pdf>
- C15. **Huiping Cao**, K. Selçuk Candan, and Maria Luisa Sapino: Skynets: Searching for Minimum Trees in Graphs with Incomparable Edge Weights. In Proc. of Intl. Conf. on Information and Knowledge Management, CIKM 2011, 1775-1784. (Acceptance rate: 15%) <http://dl.acm.org/citation.cfm?doid=2063576.2063833>
- C16. **Huiping Cao**, Shawn Bowers, Mark P. Schildhauer: Approaches for Semantically Annotating and Discovering Scientific Observational Data. In Proc. of Intl. Conf. on Database and Expert System Applications, DEXA 2011, 526-541. (Acceptance rate: 25%) <http://www.springerlink.com/content/v56047534m051171/>

- C17. Shawn Bowers, Jay Kudo, **Huiping Cao**, Mark P. Schildhauer: ObsDB: A System for Uniformly Storing and Querying Heterogeneous Observational Data. In Proc. of the IEEE Intl. Conf. on e-Science, 2010, 261-268. (Acceptance rate: 30%)
<http://www.computer.org/csdl/proceedings/escience/2010/4290/00/4290a261-abs.html>
- C18. **Huiping Cao**, Yan Qi, K. Selçuk Candan, and Maria Luisa Sapino: Feedback-driven result ranking and query refinement for exploring semi-structured data collections. In Proc. of Intl. Conf. on Extending Database Technology, EDBT 2010, 3-14. (Acceptance rate: N.A.)
<http://dl.acm.org/citation.cfm?doid=1739041.1739046>
- C19. **Huiping Cao**, Yan Qi, K. Selçuk Candan, and Maria Luisa Sapino: Exploring Path Query Results through Relevance Feedback. In Proc. of Intl. Conf. on Information and Knowledge Management, CIKM 2009, 1959-1962. (Acceptance rate: N.A.)
<http://dl.acm.org/citation.cfm?doid=1645953.1646275>
- C20. K. Selçuk Candan, **Huiping Cao**, Yan Qi, and Maria Luisa Sapino: AlphaSum: Size-Constrained Table Summarization using Value Lattices. In Proc. of Intl. Conf. on Extending Database Technology, EDBT 2009, 96-107. (Acceptance rate: 32.5%)
<http://dl.acm.org/citation.cfm?doid=1516360.1516373>
- C21. K. Selçuk Candan, **Huiping Cao**, Yan Qi, and Maria Luisa Sapino: Table Summarization with the Help of Domain Lattices. In Proc. of Intl. Conf. on Information and Knowledge Management, CIKM2008, 1473-1474. (Acceptance rate: 16%)
<http://dl.acm.org/citation.cfm?doid=1458082.1458340>
- C22. **Huiping Cao**, Nikos Mamoulis, and David W. Cheung: Discovery of Collocation Episodes in Spatiotemporal Data. In Proc. of Intl. Conf. on Data Mining, ICDM2006, 823-827. (Acceptance rate: 10%) <http://www.computer.org/csdl/proceedings/icdm/2006/2701/00/270100823-abs.html>
- C23. **Huiping Cao**, Nikos Mamoulis, and David W. Cheung: Mining Frequent Spatio-Temporal Sequential Patterns. In Proc. of Intl. Conf. on Data Mining, ICDM2005, 82-89. Received Student Travel award. (Acceptance rate: 13.8%)
<http://www.computer.org/csdl/proceedings/icdm/2005/2278/00/22780082-abs.html>
- C24. Nikos Mamoulis, **Huiping Cao**, George Kollios, Marios Hadjieleftheriou, Yufei Tao, and David W. Cheung: Mining, Indexing, and Querying Historical Spatiotemporal Data. In ACM SIGKDD Proc. of Intl. Conf. on Knowledge Discovery and Data Mining, SIGKDD2004, 236-245. (Acceptance rate: 12%) <http://dl.acm.org/citation.cfm?doid=1014052.1014080>
- C25. **Huiping Cao**, David W. Cheung, and Nikos Mamoulis: Discovering Partial Periodic Patterns in Discrete Data Sequences. In Proc. of Pacific-Asia Conf. on Knowledge Discovery and Data Mining, PAKDD2004, 653-658. (Acceptance rate: 13%)
http://link.springer.com/chapter/10.1007%2F978-3-540-24775-3_77
- C26. Yutao Shou, Nikos Mamoulis, **Huiping Cao**, Dimitris Papadias, and David W. Cheung: Evaluation of Iceberg Distance Joins. In Proc. of Intl. Symp. on Advances in Spatial and Temporal Databases, SSTD2003, 270-288. (Acceptance rate: N.A.)
http://link.springer.com/chapter/10.1007%2F978-3-540-45072-6_16

6.3 Refereed Workshop Papers

- W1. Ian Goetting**, Elisabeth Baseman, **Huiping Cao**: Causal Relationships amongst Sensors in the Trinity Supercomputer: work in progress. In Proceedings of the First Workshop on Machine Learning for Computing Systems (MLCS). Co-located with The 27th International Symposium on High-Performance Parallel and Distributed Computing (HPDC'18). <https://doi.org/10.1145/3217871.3217875>
- W2. Francisco Natividad*, Russell Y. Folk, William Yeoh, and **Huiping Cao**: On the Use of Off-the-Shelf Machine Learning Techniques to Predict Energy Demands of Power TAC Consumers. In Proc. of Intl. Workshop on Agent-Mediated Electronic Commerce and Trading Agents Design and Analysis (AMEC/TADA) 2016, pages 112-126. (Acceptance rate: N.A.)

- W3. Yifan Hao*, **Huiping Cao**, Kabi Bhattarai*, Satyajayant Misra: STK-anonymity: k-anonymity of social networks containing both structural and textual information, In Proc. of DBSocial 2013:19-24, co-located with SIGMOD 2013. (Acceptance rate: N.A.)
<http://dl.acm.org/citation.cfm?doid=2484702.2484707>
- W4. Shawn Bowers, **Huiping Cao**, Mark Schildhauer, Matt Jones, Ben Leinfelder, and Margaret O'Brien: A Semantic Annotation Framework for Retrieving and Analyzing Observational Datasets. In the third workshop on Exploiting Semantic Annotations in Information Retrieval (ESAIR2010), co-located with Intl. Conf. on Information and Knowledge Management (CIKM), 31-32, 2010. (Acceptance rate: N.A.) <http://dl.acm.org/citation.cfm?doid=1871962.1871982>

6.4 Book Chapters

- B1. **Huiping Cao**, Yan Qi, K. Selçuk Candan, and Maria Luisa Sapino: XML Data Integration: Schema Extraction and Mapping. Chapter 14 in *Advanced Applications and Structures in XML Processing*. Editors: Changqing Li and Tok Wang Ling.
- B2. Yan Qi, **Huiping Cao**, K. Selçuk Candan, and Maria Luisa Sapino: XML Data Integration: Merging, Query Processing and Conflict Resolution. Chapter 15 in *Advanced Applications and Structures in XML Processing*. Editors: Changqing Li and Tok Wang Ling.
- B3. **Huiping Cao**, Nikos Mamoulis, and David W. Cheung: Periodic Pattern Discovery from Trajectories of Moving Objects. Chapter 15 in *Geographic Data Mining and Knowledge Discovery, 2nd Edition*. Editors: Harvey J. Miller and Jiawei Han.

6.5 Posters and non-archived conference papers

- P1. Omar Navarro**, Edgar Ceh-Varela*, **Huiping Cao**: Analysis of anomaly detection algorithms on electricity consumption time series. In 2020 Emerging Researchers National (ERN) Conference in STEM, February 6-8, 2020.
- P2. Amy Worth**, **Huiping Cao**: Analysis of Traffic Accidents Using Machine Learning Techniques. In 2019 Emerging Researchers National (ERN) Conference in STEM, February 21-23, 2019.
- P3. Panika Valecha*, **Huiping Cao**, Qixu Gong*, Mai Zheng, Feng Yan, Xing Lin, Art Harkin: Analysis and Prediction of Storage Error Events for High Performance Computing Systems. In Conference on Data Analysis (CODA) 2018, March 7 - 9, 2018. Santa Fe, New Mexico.
- P4. Josue Gutierrez, Sean Flaherty**, Colby Brungard, **Huiping Cao**, David Dubois, Merrill Bean, Max Bleiweiss: Early Warning Dust Forecasting. In 29th Hispanic Engineering National Achievement Awards Conference (HENAAC) Conference, October 18-21, 2017. Pasadena, CA.
- P5. Stephanie Peña-Rivera**, **Huiping Cao**, Oswald Chong: Reducing Prediction Errors for Abnormal Energy Consumption. In CAHSI Annual Conference, co-located with HENAAC, October 18-21, 2017 in Pasadena, CA.
- P6. Brett Pelkey**, Yifan Hao*, **Huiping Cao**: Predicting Energy Consumption of Campus Buildings. In 2017 UT Dallas Undergraduate Research Expo In Computer Science, April 14, 2017. (The 1st place).
- P7. Brett Pelkey**, Stefan Ceballos**, Milan Biswal*, **Huiping Cao**, Sukumar Brahma: *Phasor Measurement Unit Monitoring Tool*, 17th Joint UTEP/NMSU Workshop on Mathematics, Computer Science, and Computational Sciences, El Paso, Texas, November 7, 2015. <http://www.cs.utep.edu/vladik/utepnmsu15.html>.
- P8. A. Al-Ghraibah, Laura E. Boucheron, R. T. James McAteer, **Huiping Cao**, Jason Jackiewicz, Bernie J. McNamara, David G. Voelz, B. Calabro, K. DeGrave, Yifan Hao, Michael Kirk, A. Pevtsov, J. McKeever, G. Taylor: *Automated Classification of Flaring Behavior in Solar Active Regions: Preliminary Results*, American Astronomical Society (AAS) Annual meeting, Austin, TX, January 10, 2012.
- P9. Yifan Hao*, **Huiping Cao**, Bernie J. McNamara, Jason Jackiewicz, R. T. James McAteer, Laura E. Boucheron, David G. Voelz,: *Intelligent Search of Solar Data*, American Astronomical Society (AAS) Annual meeting, Austin, TX, January 10, 2012.

- P10. Laura E. Boucheron, A. Al-Ghraibah, R. T. James McAteer, **Huiping Cao**, Jason Jackiewicz, Bernie J. McNamara, David G. Voelz,; *Fields, flares, and forecasts*, AAS Solar Physics Division, AAS/SPD, Las Cruces, NM, June 2011.
- P11. Bernie J. McNamara, Jason Jackiewicz, C. Lovekin, R. T. James McAteer, Laura E. Boucheron, **Huiping Cao**, David G. Voelz, Michael Kirk, G. Taylor, K. DeGrave, A. Al-Graibah, A. Pevtsov: *The Influence of Rotation on the Pulsation Spectra of B-stars*, AAS Solar Physics Division, AAS/SPD, Las Cruces, NM, June 2011.
- P12. Bernie J. McNamara, Jason Jackiewicz, R. T. James McAteer, Laura E. Boucheron, **Huiping Cao**, David G. Voelz, Michael Kirk, G. Taylor, K. DeGrave, A. Al-Graibah, C. Lovekin, A. Pevtsov: *The Pulsation Spectra of Kepler B Stars*, American Astronomical Society (AAS), May 2011.
- P13. M. Schildhauer, L. E. Bermudez, S. Bowers, P. C. Dibner, C. Gries, M. B. Jones, D. L. McGuinness, **Huiping Cao**, S. J. Cox, S. Kelling, C. Lagoze, H. Lapp, J. Madin, *A core observational data model for enhancing the interoperability of ontologically annotated environmental data*, American Geophysical Union (AGU), Fall Meeting 2010, abstract #IN44B-05. <http://adsabs.harvard.edu/abs/2010AGUFMIN44B..05S>

7 Grants

7.1 Current Grants

- G1. *Project Title*: REU Site: BIGData - Big Data Analytics for Cyber-Physical Systems.
Funding Agency and Award Number: National Science Foundation, CNS-1950121.
Amount: \$404,571
Duration: 03/01/2020 - 02/28/2023
Role: **PI**.
- G2. *Project Title*: CREST: Interdisciplinary Center for Research Excellence in Design of Intelligent Technologies for Smartgrids Phase II. (PI: Enrico Pontelli)
Funding Agency and Award Number: National Science Foundation; HDR-1914635
Amount: \$4,998,780
Duration: 02/01/2020 - 01/31/2025
Role: **Co-PI**.
- G3. *Project Title*: Preparing Highly Qualified Students with Financial Need for Careers in Computing and Cyber-Security through Evidence-Based Educational Practices.
Funding Agency and Award Number: National Science Foundation; DUE-1833630
Amount: \$ 3,969,365
Duration: 10/01/2018 - 09/30/2023
Role: **PI**.
- G4. *Project Title*: RII Track-1: The New Mexico SMART Grid Center: Sustainable, Modular, Adaptive, Resilient, and Transactive. (PI: William Michener).
Funding Agency and Award Number: National Science Foundation; OIA-1757207
Amount: \$3,933,361 (for year 1)
Duration: 09/15/2018 - 08/31/2023
Role: **Senior Personnel**; Overseeing one of the three research thrusts.
- G5. *Project Title*: BIGDATA: Collaborative Research: F: Discovering Context-Sensitive Impact in Complex Systems.
Funding Agency and Award Number: National Science Foundation; IIS-1633330
Amount: \$361,791
Duration: 09/01/2016 - 08/31/2020
Goal: The project is to design data mining and machine learning models and algorithms to discover

context-aware impacts among objects in complex systems including biological pathways, social networks, and climate systems.

Role: NMSU PI.

- G6. *Project Title:* iCREDITS: interdisciplinary Center of Research Excellence in Design of Intelligent Technologies for Smartgrids. (PI: Enrico Pontelli)
Funding Agency and Award Number: National Science Foundation; HRD-1345232
Amount: \$4,999,721 (among which \$701,655 goes to our monitoring subproject)
Duration: 02/01/2014-07/31/2020
Goal: This project is to establish the interdisciplinary Center of Research Excellence in Design of Intelligent Technologies for Smart grids (iCREDITS).
Role: **Senior Personnel;** Co-lead the subproject “Comprehensive Real Time Classification and Localization of Disturbances in Power Systems using Synchronized Measurements” with Dr. Sukumar Brahma (Co-PI; Electrical Engineering, NMSU).
Project Website: <http://icredits.nmsu.edu/>

7.2 Expired Grants

- $G_{expired1}$. *Project Title:* REU Site: BIGData - Big Data Analytics for Cyber-Physical Systems
Funding Agency and Award Number: National Science Foundation; ACI-1559723
Amount: \$359,151
Duration: 03/01/2016 - 02/28/2019
Goal: This Research Experience for Undergraduate (REU) site will support NSF’s mission to promote progress of science by introducing big data analytics in Cyber-physical Systems (CPS) to undergraduate students, helping advance the state of art, and preparing them for the future scientific workforce.
Role: **PI.**
Project Website: <http://www.cs.nmsu.edu/reu/>
- $G_{expired2}$. *Project Title:* CC*DNI Engineer: Enabling Research by Optimizing Cyberinfrastructure at NMSU and Beyond. (PI: Norma Grijalva)
Funding Agency and Award Number: National Science Foundation; ACI-1541296
Amount: \$199,931
Duration: 01/01/2016-12/31/2017
Goal: The Cyber Infrastructure Engineer (CI Engineer) accelerates scientific discovery by bridging the gap between researchers, other CI professionals, and institutional leadership.
Role: **Senior Personnel.**
- $G_{expired3}$. *Project Title:* MRI: Acquisition of an Instrument for Research in Irregularly Parallel Big Data Computation. (PI: Jonathan Cook)
Funding Agency and Award Number: National Science Foundation; CNS-1337884
Amount: \$224,074
Duration: 10/01/2013-09/30/2016
Goal: This project has acquired a computational instrument that is configured and directed to support research in irregular, mostly graph-based, computations over big data.
Role: **Co-PI.**
Project Website: <http://bigdat.nmsu.edu/>
- $G_{expired4}$. *Project Title:* Classification of Disturbance Data Generated by Phasor Measurement Units for Better Control and Protection of the Power Grid. (PI: Sukumar Brahma)
Funding Agency and Award Number: Office of the Vice President for Research, NMSU; Interdisciplinary Research Grant (IRG) 121294
Amount: \$38,490
Duration: 01/01/2013 – 12/31/2013.
Goal: The goal of this proposal is to determine the number of classes that the data stored inside the Phasor Data Concentrator (PDC) can be divided into.
Role: **Co-PI.**

- G_{expired}*5. *Project Title:* Efficient and Intelligent Management of Solar Data by Utilizing High-Performance Computing Infrastructure and Semantic Knowledge for Data-Intensive Retrieval and Exploration. (PI: Huiping Cao)
Funding Agency and Award Number: Office of the Vice President for Research, NMSU; Interdisciplinary Research Grant (IRG) 111721
Amount: \$50,000
Duration: 08/15/2011 – 08/14/2012.
Goal: The goal of the proposed research is to investigate novel science to attain more efficient and intelligent retrieval, exploration, and visualization of diverse types of the vast amount of solar data available by utilizing semantic knowledge and high-performance computing infrastructure.
Role: **PI**; Co-lead the project with Dr. Jason Jackiewicz (Co-PI; Astronomy, NMSU).

8 AWARDS

- The first recipient of the Hue and Pat McCoy Endowed Professorship in Computer Science, NMSU, 2019-2023
- Outstanding Graduate Faculty in Teaching for the year 2018-2019, the Computer Science Graduate Student Organization (CSGSO), New Mexico State University.
- Service Award, ACM Special Interest Group on Management of Data (SIGMOD) 2016

9 SERVICE

9.1 Service to the University (after joining NMSU)

- *CS department graduate advisor*, 10/15/2018 - present
- *CS department search committee chair*,
 - 10/2019 - 03/2020
 - 10/2018 - 03/2019
- *CS department service course committee chair*, 02/2013 - 01/2018
- *CS department search committee member*,
 - 11/2012 - 03/2013
 - 11/2014 - 03/2015
- *CS department publicity committee member* by maintaining social media content on Twitter, 08/2014 - 01/2018
- *Ph.D. committee member* (20) (in reverse order of the graduation time)
 Qisong Hu (Electrical engineering, 01/2020 - present), Guangwei Wen (Mechanical engineering, 01/2020 - present), Jinrui Cao (CS, 07/2019 - present), Hua Zhong (CS, 4/27/2018 - 11/1/19), Hien Nguyen (CS, 5/4/2018 - 12/6/2018), Emmanuel Utreras (CS, 12/14/2017 - present), Qianning Liu (Mathematics, 09/2016 - present), Ziwei Ma (Mathematics, 05/2016 - present), Xiaonan Zhu (Mathematics, 05/2016 - 04/09/2019), Virginia Knight (Biology, 11/2015 - 11/3/2017), Quan Do (IDoc, 11/2015 - 11/13/2019), Najah Al-shanableh (IDoc, 04/2014 - 05/2017), Raymond Paulino (Mathematics, 01/2014 - 11/4/2016), Zheng Wei (Mathematics, graduated in 05/2015), James Obert (IDoc, graduated in 05/2015), Yang Zhang (CS, graduated in 10/2015), Mohammed Sawalhah (Animal Science, graduated in 04/2013), Richard Dan McClanahan (Electrical Engineering, graduated in 04/2013), Om Dahal (Electrical Engineering, graduated in 01/2014), Ngoc-Hieu Nguyen (CS, graduated in 04/2013).
- *M.S. committee member* (27) (in reverse order of the graduation time)
 Yuan Xu (CS, 07/2018), Yuhao Lan (Biology & CS, 06/2018), Chase Gilbert (CS, 01/2018), Fredrick Ayivor (Mathematics, 8/2016) Sajal Kumar (CS, 7/2016), Frank Natividad (CS, 5/2016), Wei Wei (Statistics, 12/2015), Phillip Chen (Electrical Engineering, 05/2015), AmirSaber Sharifi (Computer

Science, 12/2014), Ruoqing Zhang (Industrial Engineering, 11/2014), Ashley Michalenko (Electrical Engineering, 07/2014), Bahar Sayoldin (Computer Science, 05/2013), Alexander Fielder (Computer Science, 05/2013), Ranjith Molgu (Computer Science, 11/2013), Karen Carter (Computer Science, 05/2013), Lakshmi Navya Makkapati (Industrial Engineering, 05/2013), Haizhou Wang (Computer Science, 04/2013), Om Dahal (Electrical Engineering, 03/2013), Alsbouaya Ali (Computer Science, 12/2012), Chaitanya Vemprala (Computer Science, 01/2012), Chandramouli Rajasekaran (Computer Science, 12/2011), John Esther (Computer Science, 11/2011), Weizhong Tian (Mathematics, 11/2011), Kabi Bhattarai (Computer Science, 11/2011), Ahed A. Elmsallati (Computer Science, 05/2011), Kueiju Kuang (Computer Science, 04/2011), Sireesha Nethikunta (Computer Science, 12/2010).

9.2 Service to the Community

Editorial Board & Journal Organizer

- Member of the editorial board of *Journal on Data Semantics* (JoDS), Publisher: Springer.
- Associate SIGMOD Information Director (August 2015 - present)

Conference and Workshop Organizer

(Note: the abbreviation names of the conferences are listed at the end of this section.)

- Co-Chair, the first workshop on Machine Learning for Computing Systems (MLCS), co-located with International Symposium on High-Performance Parallel and Distributed Computing (HPDC), Tempe, Arizona, USA, June 11, 2018. (Co-organizers: Elisabeth Baseman (Chair) and George Amvrosiadis).
- Tutorial Co-Chair, PAKDD, Singapore, May 11-14, 2020.
- Exhibits Chair, SIGMOD, San Francisco, USA. Jun 26 -Jul 1, 2016.
- Workshops Co-Chair, PAKDD, Auckland, New Zealand. Apr 19 - 22, 2016.
- Proceedings Co-Chair, SIGMOD, Melbourne, Australia. May 31- Jun 4, 2015.
- Web chair, SIGMOD, Salt Lake City, Utah, 2014.
- Web/Information Chair, SIGMOD, Scottsdale, Arizona, USA, May 20-25, 2012.

Program Committee Member

(Note: the abbreviation names of the conferences are listed at the end of this section.)

- 2020: PVLDB, AAI, SDM, ECAI, SIGIR, IJCAI, KDD, HotEdge, CIKM
- 2019: KDD, AAI, IJCAI, SDM, SIGMOD Demo, PVLDB, DASFAA, ADC
- 2018: KDD, IJCAI, CIKM, CIKM Demo, SIGMOD Demo, DASFAA, ADC
- 2017: CIKM, DASFAA, DEXA, BigCom
- 2016: KDD, SDM, CIKM, KDIR, DATA ANALYTICS
- 2015: KDD, IJCAI, CIKM, SDM, DATA ANALYTICS
- 2014: CIKM, MILCOM, DATA ANALYTICS
- 2013: DATA ANALYTICS
- 2012: ER, DATA ANALYTICS
- 2011: PAKDD
- 2010: The second IEEE Workshop on Information & Software as Services (WISS 2010), co-located with ICDE 2010
- 2009: The SIAM SDM Workshop on Multimedia Data Mining (in conjunction with SDM 2009) The first workshop on Real Time Business Intelligence (RTBI 2009), co-located with the joint International Conferences on Asia-Pacific Web Conference (APWeb) and Web-Age Information Management (WAIM 2009).
- 2008: The sixteenth Annual European Symposium on Algorithms (ESA 2018).

Journal Referee

- VLDB Journal
- IEEE Transactions on Knowledge and Data Engineering (TKDE)
- IEEE Transactions on Parallel and Distributed Systems (TPDS)
- IEEE Transactions on Semiconductor Manufacturing (TSM)
- ACM Transactions on Database Systems (TODS)
- ACM Transactions on Knowledge Discovery in Data (TKDD)
- ACM Transactions on Spatial Algorithms and Systems (TSAS)
- Knowledge and Information Systems (KAIS) Journal
- Springer Data Mining and Knowledge Discovery (DAMI)
- Distributed and Parallel Databases (DAPD)
- WWW Journal
- Journal of Applied Mathematics
- Data & Knowledge Engineering (DKE) Journal
- Geoinformatica Journal
- Book chapters for the book *Post-Mining of Association Rules – Techniques for Effective Knowledge Extraction*. Editors: Yanchang Zhao, Chengqi Zhang, Longbing Cao. Publisher: IGI Global, 2008.

Proposal Referee

- USDA NIFA panelist, year 2020
- NSF panelist, two NSF panels (IIS, CNS), year 2020
- NSF panelist, NSF IIS panels, year 2019
- NSF panelist, two NSF IIS panels, 2018
- Ad-hoc reviewer for NSF proposals, 2016
- Israel Science Foundation (ISF), 2015
- NSF panelist, one NSF IIS panel, 2015
- NSF panelist, two NSF IIS panels, 2013

Thesis Referee

- University of Turin, Italy; PhD thesis, 2019
- University of Turin, Italy; PhD thesis, 2015
- University of Melbourne, Australia; Masters of Philosophy (MPhil) thesis, 2014
- Anna university, Chennai, India; PhD thesis, 2014

Student Volunteer

- IEEE International Conference on Data Mining (ICDM), Dec., 2006.
- IEEE International Conference on Data Mining (ICDM), Nov., 2005.

Abbreviations of conference names

- KDD: ACM SIGKDD Conference on Knowledge Discovery and Data Mining
- CIKM: ACM International Conference on Information and Knowledge Management
- SDM: SAIM International Conference on Data Mining
- KDIR: International Conference on Knowledge Discovery and Information Retrieval
- IJCAI: International Joint Conference on Artificial Intelligence
- AAAI: AAAI Conference on Artificial Intelligence
- SIGMOD: ACM International Conference on Management of Data
- PVLDB: International Conference on Very Large Data Bases
- ADC: Australasian Database Conference
- DASFAA: International Conference on Database Systems for Advanced Applications
- ER: International Conference on Conceptual Modeling
- PAKDD: Pacific-Asia Conference on Knowledge Discovery and Data Mining
- MILCOM: IEEE Military Communications Conference
- DATA ANALYTICS: International Conference on Data Analytics

10 OUTREACH

- 09/2019, Led the organization of the data science workshop at CAHSI summit, which is co-located with GMiS - HENAAC Conference, September 25 - 29, 2019 Disney's Coronado Springs Resort, Lake Buena Vista, FL.
- 04/14/2016, Presented *5V's of Big Data*, NMSU Spring Salon Discovery.
- 04/15/2015, Presented *The Role of Big Data in Smartgrids: How to Guarantee Reliability and Resilience* (with Dr. Sukumar Brahma) for the Academy for Learning in Retirement, to seniors in Good Samaritan Village, Las Cruces, New Mexico.
- 08/2014 - 10/2015, Initiated and working on a memorandum agreement for an NMSUCS-SEIE dual program between the School of Economic Information Engineering (SEIE), Southwestern University of Finance and Economics, and the Department of Computer Science (CS), College of Arts and Sciences, New Mexico State University.
- 05/2013, Visited two Chinese universities, Southwestern University of Finance and Economics and Nanjing University of Posts and Telecommunications, to enable student recruitment, collaboration, and faculty/student exchanges.

11 PRESENTATIONS

- 4/26/2019: CASCADE Workshop/Retreat on Big Data Challenges, Techniques, and Applications, Arizona State University, Tempe, AZ.
Big data @ NMSU.
- 4/26/2019: CASCADE Workshop/Retreat on Big Data Challenges, Techniques, and Applications, Arizona State University, Tempe, AZ.
Data-driven Decision Making in Smart Grids.
- 4/17/2019: Sandia National Lab A4H Field Day, Sandia National Lab, Albuquerque, NM.
Data-driven decision making.
- 1/29/2019, New Mexico Research Spotlight Forum Artificial Intelligence & Machine Learning, Sandia National Lab, Albuquerque, NM. *Knowledge discovery and data mining from time series and graph data.*

- 10/18/2017, Dept. of Computer Science, University of New Mexico (UNM), Albuquerque, NM.
Designing and Learning Probabilistic Graphical Models for Graph Mining.
- 04/17/2017, Information School, Renmin University of China (RUC), Beijing, China
Influence Discovery from Graph-structured Data.
- 11/18/2016, Business School, NMSU
Techniques for Big Data Analysis.
- 09/23/2016, Dept. of Computer Science, New Mexico Institute of Mining and Technology (NMT), Socorro, NM
Aspect-level Influence Discovery from Graphs.
- 04/04/2015, 5th Southwest Energy Science and Engineering Symposium.
Interdisciplinary Center of Research Excellence in Design of Intelligent Technologies for Smartgrids: Research Agenda (with Dr. Sukumar Brahma).
- 11/21/2014, Department of Computer Science, University of Texas at El Paso (UTEP).
Detecting Influence Relationships from Graphs.
- 10/04/2013, Annual SACNAS National conference.
Database Support for Enabling Data-Discovery Queries Over Semantically-Annotated Observational Data.
- 05/30/2013, Department of Computer Science & Technology, Nanjing University of Posts and Telecommunications.
STK-Anonymity: K-anonymity of Social Networks Containing both Structural and Textual Information.
- 01/27/2011, Department of Mathematical Sciences, New Mexico State University (NMSU).
Feedback-driven Result Ranking and Query Refinement for Exploring Semi-structured Data Collections.
- 11/12/2010, Department of Computer Science, University of Texas at El Paso (UTEP).
Feedback-driven Result Ranking and Query Refinement for Exploring Semi-structured Data Collections.

12 PROFESSIONAL SOCIETY MEMBERSHIP

- Association of Computing Machinery (ACM)