

Jessica Lin

Computer Science
George Mason University
4400 University Drive, MS 4A5
Fairfax, VA 22030

Office: (703) 993-4693
Fax: (703) 993-1710
Email: jessica@cs.gmu.edu
Website: <http://www.cs.gmu.edu/~jessica>

Education

- 2005 PhD in Computer Science
University of California at Riverside
Thesis: Discovering Unusual and Non-trivial Patterns in Time Series Databases
- 2002 M.S. in Computer Science
University of California at Riverside
Project: Dimensionality Reduction by Random Projection and Latent Semantic Indexing
- 1999 B.S. in Computer Science
University of California at Riverside

Work Experience

- August 2008 – Present Assistant Professor
Department of Computer Science (department merged with the ISE department on January 1, 2008), George Mason University
- August 2005 – 2007 Assistant Professor
Department of Information and Software Engineering, George Mason University
- September 1999 – June 2005 Research and Teaching Assistant
Department of Computer Science, University of California at Riverside
- June 2004 – September 2004 Research Intern
AT&T Research, Information Visualization Group
Design and development of visualization software for epidemic outbreak detection
- June 2003 – September 2003 Research Intern
Microsoft Research, Machine Learning and Applied Statistics Group
Design and development of time series forecasting techniques

Publications

Edited Volumes As Editor

1. Guido Cervone, **Jessica Lin**, and Nigel Waters (Eds.). 2012. *Spatio-Temporal Data Mining for Geoinformatics: Methods and Applications*. Springer-Verlag. To be released in 2012.
2. **Jessica Lin**, Guido Cervone, and Nigel Waters (Eds.). 2010. *Proceedings of the 1st ACM SIGSPATIAL International Workshop on Data Mining for Geoinformatics*. ACM, New York, NY, USA.

Refereed Journal Articles

3. **Jessica Lin**, Rohan Khade, and Yuan Li. 2012. Rotation-invariant similarity in time series using Bag-of-Patterns representation. *Journal of Intelligent Information Systems*. To Appear.
4. Chun-Kit Ngan, Alexander Brodsky, and **Jessica Lin**. 2012. Multi-event decision making over multivariate time series. *International Journal of Information and Decision Sciences*. To Appear.
5. **Jessica Lin**, Eamonn Keogh, Li Wei, and Stefano Lonardi. 2007. Experiencing SAX: a novel symbolic representation of time series. *Data Mining and Knowledge Discovery Journal (DAMI)*, 15(2): 107-144.
6. Stefano Lonardi, **Jessica Lin**, Eamonn Keogh, and Bill Chiu. 2007. Efficient discovery of unusual patterns in time series. *New Generation Computing*, 25(1): 61-93.
7. Eamonn Keogh, **Jessica Lin**, Sang-Hee Lee, and Helga Van Herle. 2006. Finding the most unusual time series subsequence: algorithms and applications. *Knowledge and Information Systems Journal (KAIS)*, 11(1): 1-27.
8. Eamonn Keogh, **Jessica Lin**, Ada Fu & Helga Van Herie. 2006. Finding unusual medical time series subsequences: algorithms and applications. *IEEE Transactions on Information Technology in Biomedicine*, 10(3): 429-439.
9. **Jessica Lin**, Eamonn Keogh, and Stefano Lonardi. 2005. Visualizing and discovering non-trivial patterns in large time series databases. *Information Visualization*, 4(2): 61-82.
10. Eamonn Keogh and **Jessica Lin**. 2005. Clustering of time series subsequences is meaningless: implications for previous and future research. *Knowledge and Information Systems Journal (KAIS)*, 8(2): 154-177.
11. **Jessica Lin**, Eamonn Keogh, 2004, Finding or not finding rules in time series. *Applications of Artificial Intelligence in Finance and Economics (Advances in Econometrics, Volume 19)*, pp.175-201.

Articles in Refereed Conference Proceedings

12. Yuan Li, **Jessica Lin**, and Tim Oates. 2012. Visualizing variable-length time series motifs. In *Proceedings of the 2012 SIAM International Conference on Data Mining*. ♦ Acceptance Rate: 27%. To Appear.
13. Chun-Kit Ngan, Alexander Brodsky, and **Jessica Lin**. 2011. A service framework for learning, querying, and monitoring multivariate time series. In *Proceedings of the 13th International Conference on Enterprise Information Systems*. Pages 92-101. ♦ **Best Student Paper Award** ♦ Full paper Acceptance Rate: 14%
14. **Jessica Lin** and Yuan Li. 2010. Finding approximate frequent patterns in streaming medical data. In *Proceedings of the 23rd IEEE International Symposium on Computer-Based Medical Systems*. IEEE Computer Society, Washington DC, USA.
15. Chun-Kit Ngan, Alexander Brodsky, and **Jessica Lin**. 2010. Decisions on multivariate time series:

- combining domain knowledge with utility maximization. In *Supplemental Proceedings of the 15th International Conference on Decision Support Systems*. July 7-10. Lisbon, Portugal.
16. **Jessica Lin** and Yuan Li. 2009. Finding structural similarity in time series data using Bag-of-Patterns representation. In *Proceedings of the 21st International Conference on Scientific and Statistical Database Management (SSDBM 2009)*, Marianne Winslett (Ed.). Springer-Verlag, Berlin, Heidelberg, Pages 461-477. ♦ Full Paper Acceptance rate: 38%
 17. **Jessica Lin** and Yuan Li. 2009. Finding structurally different medical data. In *Proceedings of the 22nd IEEE International Symposium on Computer-Based Medical Systems*. IEEE Computer Society, Washington DC, USA. Pages 1-8.
 18. **Jessica Lin** and David Etter. 2008. Exact and approximate reverse nearest neighbor search in multimedia data. In *Proceedings of the SIAM International Conference on Data Mining*. Pages 656-667. ♦ Short Paper. Overall Acceptance Rate 27%: 14% Full + 13% Short
 19. **Jessica Lin** and Eamonn Keogh. 2006. Group SAX: Extending the notion of contrast sets to time series and multimedia data. In *Proceedings of the 10th European Conference on Principles and Practice of Knowledge Discovery in Databases*. Berlin, Germany. Sept 18-22. Pages 284-296. Lecture Notes in Computer Science, Springer. ♦ Full Paper Acceptance Rate: 14.5%
 20. Eamonn Keogh, **Jessica Lin**, and Ada Fu. 2005. HOT SAX: Efficiently finding the most unusual time series subsequence. In *Proceedings of the 5th IEEE International Conference on Data Mining (ICDM)*. Nov 27-30. Houston, TX. Pages 226-233. IEEE Computer Society. ♦ Full Paper Acceptance Rate: 10.9%
 21. **Jessica Lin**, Eamonn Keogh, Ada Fu, and Helga Van Herie. 2005. Approximations to magic: finding unusual medical time series. In *Proceedings of the 18th International Symposium on Computer-Based Medical Systems*. IEEE Computer Society, Washington DC, USA. Pages 329-334. ♦ Full Paper Acceptance Rate: 36.4%
 22. **Jessica Lin**, Michail Vlachos, Eamonn Keogh, Dimitrios Gunopulos, Jianwei Liu, Shoujian Yu, and Jiajin Le. 2005. A MPAA-based iterative clustering algorithm augmented by nearest neighbors search for time-series data streams. In *Proceedings of the 9th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-05)*. Lecture Notes in Computer Science, Springer. Pages 333-342. ♦ Full Paper Acceptance Rate: 29.8%
 23. **Jessica Lin**, Eamonn Keogh, Stefano Lonardi, Jeffrey P. Lankford, and Donna M. Nystrom. 2004. Visually mining and monitoring massive time series. In *Proceedings of the tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '04)*. ACM, New York, NY, USA, 460-469. ♦ Full Paper Acceptance Rate: 30%
 24. **Jessica Lin**, Eamonn Keogh, Stefano Lonardi, Jeffrey P. Lankford, and Daonna M. Nystrom. 2004. VizTree: a tool for visually mining and monitoring massive time series databases. In *Proceedings of the 30th International Conference on Very large Data Bases - Volume 30 (VLDB '04)*. VLDB Endowment. Pages 1269-1272.
 25. **Jessica Lin**, Michail Vlachos, Eamonn Keogh, and Dimitrios Gunopulos. 2004. Iterative incremental clustering of time series. In *Proceedings of the IX Conference on Extending Database Technology (EDBT)*. Lecture Notes in Computer Science, Springer. Pages 106-122. ♦ Full Paper Acceptance: 14.3%
 26. Eamonn Keogh, **Jessica Lin**, and Wagner Truppel. 2003. Clustering of time series subsequences is meaningless: implications for past and future research. In *Proceedings of the 3rd IEEE International Conference on Data Mining (ICDM)*. IEEE Computer Society. Pages 115-122. ♦ Full Paper Acceptance Rate 11.6%
 27. **Jessica Lin**, Eamonn Keogh, and Wagner Truppel. 2003. (Not) Finding rules in time series: a surprising result with implications for previous and future research. In *Proceedings of the 2003 International Conference on Artificial Intelligence*. Las Vegas, NV. June 23-26. Pages 55-61.

28. Pranav Patel, Eamonn Keogh, **Jessica Lin**, and Stefano Lonardi. 2002. Mining motifs in massive time series databases. In *Proceedings of the 2002 IEEE International Conference on Data Mining (ICDM '02)*. IEEE Computer Society, Washington, DC, USA, 370-377. ♦ Full Paper Acceptance Rate: 19.5%

Refereed Workshop Papers

29. Chun-Kit Ngan, Alexander Brodsky, and **Jessica Lin**. 2011. Multi-event decision making over multivariate time series. In *Proceedings of EWG-DSS London 2011 Workshop on Decision Systems*. Birbeck, UK. June 23-34.
30. **Jessica Lin**, Guido Cervone, and Pasquale Franzese. 2010. Assessment in error in air quality models using dynamic time warping. In *Proceedings of the 1st International Workshop on Data Mining for Geoinformatics, in conjunction with SIGSPATIAL GIS 2010*. San Jose, CA. Nov 2, 2010. Pages 38-44.
31. Yuan Li and **Jessica Lin**. 2010. Approximate variable-length time series motif discovery using grammar inference. In *Proceedings of the 10th International Workshop on Multimedia Data Mining, in conjunction with SIGKDD 2010*. ACM, New York, NY, USA. Pages 1-9.
32. Eiman Al-Shammari and **Jessica Lin**. 2008. A new Arabic stemming algorithm. In *Proceedings of the 2008 ISCA Workshop on Experimental Linguistics*. Athens, Greece. Aug 25-27.
33. Eiman Al-Shammari, **Jessica Lin**. 2008. A novel Arabic lemmatization algorithm. In *Proceedings of the 2nd SIGIR Workshop on Analytics for Noisy Unstructured Text Data*. Singapore, July 24-27, 2008. p. 113-118.
34. Eiman Al-Shammari and **Jessica Lin**. 2008. Towards an error-free Arabic stemming. In *Proceeding of the 2nd ACM workshop on Improving non English web searching (INEWS), at the 17th ACM Conference on Information and Knowledge Management*. ACM, New York, NY, USA. Pages 9-16. ♦ Full Paper Acceptance Rate: 37%
35. David DeBarr and **Jessica Lin**. 2007. Time series classification challenge experiments. In *Proceedings of the Workshop and Challenge on Time Series Classification, at the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Jose, CA. Aug 12-15.
36. **Jessica Lin**, Vlachos, M, Keogh, E., & Gunopulos, D. 2003. Multi-resolution k-means clustering of time series and application to images. In *Proceedings of the 4th SIGKDD Workshop on Multimedia Data Mining, in conjunction with SIGKDD 2003*. 10 pages.
37. **Jessica Lin**, Eamonn Keogh, Stefano Lonardi, and Bill Chiu. 2003. A symbolic representation of time series, with implication for streaming algorithms. In *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD '03)*. ACM, New York, NY, USA. Pages 2-11. ♦ Full Paper Acceptance Rate: 28.5%
38. **Jessica Lin**, Eamonn Keogh, and Wagner Truppel. 2003. Clustering of streaming time series is meaningless. In *Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery (DMKD '03)*. ACM, New York, NY, USA. Pages 2-11. ♦ Full Paper Acceptance Rate: 28.5%
39. Michail Vlachos, **Jessica Lin**, Eamonn Keogh, and Dimitrios Gunopulos. 2003. A wavelet-based anytime algorithm for k-means clustering of time series. In *Proceedings of the Workshop on Clustering High Dimensional Data and Its Applications, at the 3rd SIAM International Conference on Data Mining*. San Francisco, CA. May 3, 2003. 12 pages.
40. **Jessica Lin** and Gunopulos, D. 2003. Dimensionality reduction by random projection and latent semantic indexing. In *Proceedings of the Text Mining Workshop, at the 3rd SIAM International Conference on Data Mining*. San Francisco, CA. May 3, 2003. 10 pages.
41. **Jessica Lin**, Eamonn Keogh, Pranav Patel, and Stefano Lonardi. 2002. Finding motifs in time series. In *Proceedings of the 2nd Workshop on Temporal Data Mining, at the 8th ACM SIGKDD International*

Conference on Knowledge Discovery and Data Mining. Edmonton, Alberta, Canada. July 23-26, 2002. 11 pages.

Invited Paper

42. Ada Fu, Oscar Leung, Eamonn Keogh, and **Jessica Lin**. 2006. Finding time series discords based on Haar transform. In *Proceedings of the 2nd International Conference on Advanced Data Mining and Applications*. Xi'an, China. Aug 14-18. Pages 31-41.

Refereed Chapters in Books

43. **Jessica Lin**, Sheri Williamson, Kirk Borne, and David DeBarr. 2012. Pattern recognition in time series. *Advances in Machine Learning and Data Mining for Astronomy*. Eds. Kamal, A., Srivastava, A., Way, M., and Scargle, J. Chapman & Hall. *To Appear*.
44. Chun-Kit Ngan, Alexander Brodsky, and **Jessica Lin**. 2011. An event-based service framework for learning, querying, and monitoring multivariate time series. *Lecture Notes in Business Information Processing*. Springer-Verlag. *To Appear*.
45. Chotirat Ann Ratanamahatana, **Jessica Lin**, Dimitrios Gunopulos, Eamonn Keogh, Michail Vlachos, and Gautam Das. 2010. Mining time series data. *Data Mining and Knowledge Discovery Handbook 2010*, 2nd Edition. Eds. Oded Maimon, Lior Rokach. Springer. Pages 1049-1077
46. **Jessica Lin**, Michail Vlachos, Eamonn Keogh, and Dimitrios Gunopulos. 2007. Multi-resolution time series clustering and application to images. *Multimedia Data Mining and Knowledge Discovery*, Eds. Valery Al Petrushin and Latifur Khan. Springer. Pages 58-79.
47. Chotirat Ann Ratanamahatana, **Jessica Lin**, Dimitrios Gunopulos, Eamonn Keogh, Michail Vlachos, and Gautam Das. 2005. Mining time series data. *Data Mining and Knowledge Discovery Handbook 2005*. Eds. Oded Maimon, Lior Rokach. Springer. Pages 1069-1103.

Other Publications

48. Guido Cervone, **Jessica Lin**, Pasquale Franzese. 2011. Addressing wind direction uncertainty in source estimation through dynamic time warping, In *Proceedings of the 91st American Meteorological Society Annual Meeting, Computational Intelligence Methods and Their Applications to Environmental Science*, Seattle, WA, January 2011.
49. **Jessica Lin**, Guido Cervone, and Nigel Waters. 2011. DMGI 2010 workshop report: The First ACM SIGSPATIAL international Workshop on Data Mining for Geoinformatics. SIGSPATIAL Special 3(1): 6-7.
50. Eiman Al-Shammari and **Jessica Lin**. 2008. Automated Corpora Creation Using a Novel Arabic Stemming Algorithm. In *proceedings of the 2008 International Symposium on Using Corpora in Contrastive and Translation Studies (UCCTS)*. Hanzhou, China. Sept 25-27.
51. Eamonn Keogh, **Jessica Lin**, Stefano Lonardi, and Bill Chiu. 2004. We have seen the future, and it is symbolic. In *Proceedings of the Second Workshop on Australasian Information Security, Data Mining and Web Intelligence, and Software Internationalisation - Volume 32 (ACSW Frontiers '04)*, J. Hogan, P. Montague, M. Purvis, and C. Steketee (Eds.), Vol. 32. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, Page 83.

Dissertation

- **Jessica Lin**. 2005. *Discovering Unusual and Non-Trivial Patterns in Massive Time Series Databases*. PhD thesis. Department of Computer Science, University of California at Riverside.

Software

- VizTree (demo can be found at: <http://www.cs.gmu.edu/~jessica/viztree.htm>)
- SAX (Symbolic Aggregate approXimation): <http://www.cs.gmu.edu/~jessica/sax.htm>
- Danger of Dimensionality Reduction: <http://www.cs.gmu.edu/~jessica/DimReducDanger.htm>

Invited Talks and Presentations

- 2009 George Mason University, Guest Speaker in EOS 753 (Observing the Earth and Its Climate),
Department of Geography and Geoinformation Science
Data Mining in Time Series Databases
- 2009 Virginia Tech, Department of Computer Science
Finding Non-Trivial Patterns and Structural Similarity in Time Series Databases
- 2007 USPS Office of Inspector General
Time Series Anomaly Detection
- 2007 Academia Sinica, Taiwan, Institute of Information Science
Discovering Unusual and Non-trivial Patterns in Massive Time Series
- 2007 National Chengchi University, Taiwan, Department of Computer Science
Discovering Unusual and Non-trivial Patterns in Massive Time Series

Professional Service

Workshop Chair

- SIGSPATIAL International Workshop on Data Mining for Geoinformatics, in cooperation with SIGSPATIAL GIS 2010.

Workshop Program Chair

- Special Session on Data Mining for Geoinformatics in Cooperation with the 17th International Conference on Geoinformatics, 2009.
- International Workshop on Spatial and Spatiotemporal Data Mining (SSTD 09) in Cooperation with IEEE ICDM 2009
- International Workshop on Spatial and Spatiotemporal Data Mining (SSTD 08) in Cooperation with IEEE ICDM 2008

Program Committee Member

2011

- International Conference on Data Mining (ICDM)
- ACM International Conference on Knowledge Discovery and Data Mining (KDD)
- European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECML/PKDD)
- Data-Driven Decision Guidance and Support Systems Workshop, in conjunction with the 28th IEEE International Conference on Data Engineering (ICDE)

2010

- International Conference on Data Mining (ICDM)
- European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECML/PKDD)
- International Conference on Machine Learning and Applications (ICMLA)

2009

- International Conference on Data Mining (ICDM)
- European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECML/PKDD)

2008

- International Conference on Data Mining (ICDM)
- European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECML/PKDD)

2007

- International Conference on Data Mining (ICDM)
- ACM International Conference on Knowledge Discovery and Data Mining (KDD)
- European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases (ECML/PKDD)

2006

- ACM International Conference on Knowledge Discovery and Data Mining (KDD)

Journal Reviewer

2011

- Transaction on Knowledge and Data Engineering (TKDE)
- Journal of Intelligent Information Systems (JIIS)
- ACM Computing Survey

2010

- Transaction on Knowledge and Data Engineering (TKDE)
- Data Mining and Knowledge Discovery Journal (DAMI)

2009

- Transaction on Knowledge and Data Engineering (TKDE)
- Knowledge and Information Systems Journal (KAIS)
- ACM Computing Survey

2008

- Transaction on Knowledge and Data Engineering (TKDE)
- Knowledge and Information Systems Journal (KAIS)
- Data Mining and Knowledge Discovery Journal (DAMI)
- Journal of Classification

2007

- Transaction on Knowledge and Data Engineering (TKDE)
- Knowledge and Information Systems Journal (KAIS)
- International Journal on Very Large Data Bases (VLDB Journal)

2006

- Transaction on Knowledge and Data Engineering (TKDE)
- Knowledge and Information Systems Journal (KAIS)
- Data Mining and Knowledge Discovery Journal (DAMI)
- Journal of Intelligent Information Systems (JIIS)