

# Jessica H. Lin

---

CONTACT INFORMATION	Computer Science George Mason University 4400 University Drive, MS 4A5 Fairfax, VA 22030	Office: (703) 993-4693 Fax: (703) 993-1638 Email: <a href="mailto:jessica@cs.gmu.edu">jessica@cs.gmu.edu</a> Website: <a href="http://www.cs.gmu.edu/~jessica">http://www.cs.gmu.edu/~jessica</a>
RESEARCH INTERESTS	Data mining and knowledge discovery, particularly on time series data, including indexing, clustering, motif discovery, anomaly detection, rule discovery, and visualization. Text mining.	
EDUCATION	<b>Ph.D., Computer Science (Research: Database/Data Mining)</b> University of California, Riverside Dissertation Title: Discovering Unusual and Non-trivial Patterns in Time Series Databases Advisor: Dr. Eamonn Keogh June 2005	
	<b>M.S., Computer Science (Research: Database/Data Mining)</b> University of California, Riverside Project: Dimensionality Reduction by Random Projection and Latent Semantic Indexing Advisor: Dr. Dimitrios Gunopulos Mar 2002	
	<b>B. S., Computer Science</b> University of California, Riverside July, 1999	
TEACHING	<b>Data Structures</b> (Fall 2005, Spring 2005, Fall 2006) <b>Database Management Systems</b> (Fall 2007, Spring 2008, Fall 2008) <b>Data Mining for Multimedia Data</b> (Fall 2008) <b>Special Topics in Data Mining</b> (Fall 2006, Spring 2007, Fall 2007)	
PUBLICATIONS	<b>Journal Articles</b> <ol style="list-style-type: none"><li>1. Lin, J., Keogh, E., Li, W. &amp; Lonardi, S. (2007). <b>Experiencing SAX: A Novel Symbolic Representation of Time Series</b>. <i>Data Mining and Knowledge Discovery Journal</i>. To Appear.</li><li>2. Lonardi, S., Lin, J., Keogh, E. &amp; Chiu, B. (2006). <b>Efficient Discovery of Unusual Patterns in Time Series</b>. <i>Special Issue of New Generation Computing Journal</i>.</li><li>3. Keogh, E., Lin, J. &amp; Fu, A. (2006). <b>Finding the Most Unusual Time Series Subsequence: Algorithms and Applications</b>. <i>Knowledge and Information Systems (KAIS)</i>. Springer-Verlag.</li><li>4. Keogh, E., Lin, J., Fu, A. &amp; Van Herie, H. (2005). <b>Finding the Unusual Medical Time Series: Algorithms and Applications</b>. <i>IEEE Transactions on Information Technology in Biomedicine</i>.</li><li>5. Lin, J., Keogh, E. &amp; Lonardi, S. (2005). <b>Visualizing and Finding Patterns in Time Series</b>. <i>Information Visualization Journal</i>.</li><li>6. Keogh, E., Lin, J. &amp; Truppel, W. (2004). <b>Clustering of Time Series Subsequences is Meaningless: Implications for Past and Future Research</b>. <i>Knowledge and Information Systems (KAIS)</i>, Springer-Verlag.</li><li>7. Lin, J. &amp; Keogh, E. (2004). <b>Finding or Not Finding Rules in Time Series</b>. <i>Special Issue of Advances in Econometrics</i>.</li></ol>	

### Conference Proceedings

8. Al-Shammari, E. & Lin, J. (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. To Appear.
9. Lin, J. & Etter, D. (2008). **Exact and Approximate Reverse Nearest Neighbor Search in Multimedia Data**. *SIAM International Conference on Data Mining*.
10. Lin, J. & Keogh, E. (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.
11. Fu, A., Leung, O., Keogh, E. & Lin, J. (2006). **Finding Time Series Discords Based on Haar Transform**. In *proceedings of the 2<sup>nd</sup> International Conference on Advanced Data Mining and Applications*. Xi'an, China. Aug 14-18.
12. Keogh, E., Lin, J. & Fu, A. (2005). **HOT SAX: Efficiently Finding the Most Unusual Time Series Subsequence**. In *proceedings of the 5<sup>th</sup> IEEE International Conference on Data Mining (ICDM)*. Nov 27-30. Houston, TX.
13. Lin, J., Keogh, E., Fu, A., & van Herie, H. (2005). **Approximations to Magic: Finding Unusual Medical Time Series**. In *proceedings of the 18<sup>th</sup> International Symposium on Computer-Based Medical Systems*. June 23-24. Dublin, Ireland.
14. Lin, J., Keogh, E., Lonardi, S., Lankford, J. & Nystrom, D. (2004). **Visually Mining and Monitoring Massive Time Series**. In *proceedings of the International Conference on Knowledge Discovery and Data Mining (SIGKDD)*. Aug 22-25. Seattle, WA.
15. Lin, J., Keogh, E., Lonardi, S., Lankford, J. & Nystrom, D. (2004). **VizTree: a Tool for Visually Mining and Monitoring Massive Time Series**. In *proceedings of the 30th International Conference on Very Large Databases (VLDB)*. Aug 29-Sep 3. Toronto, CA.
16. Lin, J., Vlachos, M., Keogh, E., & Gunopulos, D. (2004). **Iterative Incremental Clustering of Time Series**. *IX Conference on Extending Database Technology (EDBT)*. March 14-18, 2004.
17. Keogh, E., Lin, J. & Truppel, W. (2003). **Clustering of Time Series Subsequences is Meaningless: Implications for Past and Future Research**. In *proceedings of the 3<sup>rd</sup> IEEE International Conference on Data Mining (ICDM)*. Melbourne, FL. Nov 19-22, 2003.
18. Lin, J. Keogh, E. & Truppel, W. (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, Special Session on Applications of AI in Finance and Economics*. Las Vegas, NV. June 23-26, 2003.
19. Patel, P., Keogh, E., Lin, J., & Lonardi, S. (2002). **Mining Motifs in Massive Time Series Databases**. In *proceedings of the 2002 IEEE International Conference on Data Mining (ICDM)*. Maebashi City, Japan. Dec 9-12, 2002.

### Book Chapters

20. Lin, J., Vlachos, M., Keogh, E. & Gunopulos, D. (2006). **Multi-Resolution Time Series Clustering and Application to Images**. *Multimedia Data Mining and Knowledge Discovery*, Eds. Valery Al Petrushin and Latifur Khan.

Springer.

21. Ratanamahatana, C. A., Lin, J., Gunopulos, D., Keogh, E. & Vlachos, M. (2005). **Mining Time Series Data**. *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*. Kluwer Academic Publishers.

### Workshops

22. Al-Shammari, E. & Lin, J. (2008). **A New Arabic Stemming Algorithm**. In proceedings of the 2008 ISCA Workshop on Experimental Linguistics. Athens, Greece. Aug 25-27. To Appear.
23. Al-Shammari, E. & Lin, J. (2008). **A Novel Arabic Lemmatization Algorithm**. In proceedings of the 2<sup>nd</sup> SIGIR Workshop on Analytics for Noisy Unstructured Text Data. Singapore, July 24-27, 2008. p. 113-118.
24. DeBarr, D. & Lin, J. (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.
25. Lin, J., Vlachos, M., Keogh, E., & Gunopulos, D. (2003). **Multi-Resolution K-Means Clustering of Time Series and Application to Images**. In *proceedings of the 4<sup>th</sup> SIGKDD Workshop on Multimedia Data Mining*. Washington D.C., Aug 24-27, 2003.
26. Lin, J., Keogh, E., Lonardi, S. & Chiu, B. (2003). **A Symbolic Representation of Time Series, with Implication for Streaming Algorithms**. In *proceedings of the 8<sup>th</sup> ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*. San Diego, CA. June 13, 2003. <http://www.ise.gmu.edu/~jessica/sax.htm>
27. Lin, J. Keogh, E. & Truppel, W. (2003). **Clustering of Streaming Time Series is Meaningless**. In *proceedings of the 8<sup>th</sup> ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery*. San Diego, CA. June 13, 2003. (Acceptance rate: 28%).
28. Lin, J., Vlachos, M., Keogh, E. & Gunopulos, D. (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 3<sup>rd</sup> SIAM International Conference on Data Mining*. San Francisco, CA. May 3, 2003.
29. Lin, J. & Gunopulos, D. (2003). **Dimensionality Reduction by Random Projection and Latent Semantic Indexing**. In *proceedings of the Text Mining Workshop, at the 3<sup>rd</sup> SIAM International Conference on Data Mining*. San Francisco, CA. May 3, 2003.
30. Lin, J., Keogh, E., Patel, P. & Lonardi, S. (2002). **Finding Motifs in Time Series**. In the *2<sup>nd</sup> Workshop on Temporal Data Mining, at the 8<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Edmonton, Alberta, Canada. July 23-26, 2002.

PROFESSIONAL  
EXPERIENCE

**Assistant Professor**  
Computer Science Department

8/05 – Present

**Research Intern**  
AT&T Research, Information Visualization  
[north@research.att.com](mailto:north@research.att.com)

6/04 – 9/04

**Research Intern** 6/03 – 9/03  
Microsoft Research, Machine Learning and Applied Statistics.

**Summer Intern** 6/01 – 9/01  
Environmental Services Research Institute, Redlands, California

**Programmer** 6/98 – 6/00  
George E. Brown Jr. Salinity Laboratory,  
U.S. Department of Agriculture (U.S.D.A.)

PROFESSIONAL  
ACTIVITIES

- Program Chair, International Workshop on Spatial and Spatiotemporal Data Mining (SSTDM 08) in Cooperation with IEEE ICDM 2008
- Program Committee, PKDD/ECML 2007, 2008
- Program Committee, ICDM 2007, 2008
- Program Committee, KDD 2006, 2007
- Reviewer, KAIS
- Reviewer, TKDE
- Reviewer, DAMI (Data Mining and Knowledge Discovery Journal)
- Reviewer, The Visual Computer Journal
- Reviewer, *Special Issue of Advances in Econometrics - Applications of AI in Finance & Econ*
- External Reviewer: ICDE 2006, CBMS 2005, TDM 2005, VLDB 2004, DMWI 2003, ICML 2003, CIKM 2002.