

# Xiaoyue Elaine Wang

## PhD Candidate

Data Mining Lab  
xwang@cs.ucr.edu  
www.cs.ucr.edu/~xwang/

## Campus Address

Department of CSE  
Univ. of California Riverside  
Riverside, CA 92521

## OBJECTIVE

Software engineering position with special interest in the area of data mining, machine learning

## EDUCATION

- 09/2006- present **Ph.D.** (Expected 12/2009)  
University of California, Riverside, USA  
Major: Computer Science, Advisor: Prof. Eamonn Keogh, GPA: 3.9/4.0
- 09/2002- 06/2006 **B.S.**  
Wuhan University, Wuhan, China  
Major: Computer Science, GPA: 89.1/100

## RESEARCH INTERESTS

- Data mining/Machine learning techniques with applications to time series data
- Image retrieving and analysis
- Shape indexing/matching problems
- Social network analysis

## PROFESSIONAL EXPERIENCE

**University of California Riverside** Riverside CA  
Research Assistant

07/2007 - present

- **Similarity Measures for Time Series Data**
  - Introduced an advanced version of Dynamic Time Warping algorithm which can improve the classification accuracy when the time series data contains a certain amount of noises on both ends.
  - Produced an augmented Euclidean distance measure for time series of shapes.
- **Multi-feature Image Matching**
  - Proposed an efficient method to combine different similarity measures such as shape and color for image matching.
- **Indexing Techniques with Constraint on Space and Time**
  - Introduced a **just-in-time** indexing technique to speed up the joins in massive real valued datasets, showed applications in domains as diverse as sensor mining, blog indexing and historical manuscript mining.
  - Implemented an **anyspace** indexing algorithm with applications to data mining.
  - Showed a framework working in the scenario of **in memory** data, e.g. sensor data, based on the time series bitmaps technique. Efficient classifiers in this case can be updated in constant time and space in the face of very high data arrival rates.
- **Semi-supervised Learning Algorithms on Shape Matching**
  - Produced a more accurate semi-supervised learning algorithm which leveraged off a novel observation about the effects of shape complexity on distance measures. Showed an effective way to determine the complexity of a shape in the time series representation.
- **Demos**
  - Built a demo website on content-based image indexing algorithms for a large database of scientific nematode images. It allows users to retrieve different information from the database, e.g. to find most similar species for the given query nematode image. The website is built and maintained using HTML, PostgreSQL,

PHP and JavaScript.

-Created a **GUI** tool in Matlab that converts photos to the appropriate representation. Now the tool is used to index petroglyph images for Department of Anthropology in UCR.

**University of California Riverside** Riverside CA  
Teaching Assistant

09/2006 - 06/2007

- Assisted students with computer problems, troubleshooted hardware and software problems.

**Sprint Applied Technology Lab** Burlingame, CA  
Research Intern

06/2008 - 09/2008

Social network analysis

- Considered the problem of assigning weights to edges in a **social graph** such that the edges accurately reflect the relationship between nodes. Compared various edge weighting techniques, ranging from simple aggregate-volume metrics to time-series analysis of node interaction data for the **node label prediction problem**.

## SKILLS

- Programming: C/C++, BASH scripting, PHP, HTML, Postgresql, Latex
- Software: Matlab
- Operating System: Mac OS X, Linux, Windows NT/2000/XP

## AWARDS

- Dean's Distinguished Fellowship Award, University of California, Riverside, 2006-2008
- Teng Fei Special Scholarship Award, Wuhan University, China, 2005-2006

## JOURNAL

- **Wang, X.**, Ye, L., Keogh, E and C. Shelton. *Annotating Historical Archives of Images*. (International Journal of Digital Library Systems, to appear January 2010)

## SELECTED PUBLICATIONS

- **Wang, X.**, Ye, L., Keogh, E and C. Shelton. *Annotating Historical Archives of Images*. (**runner up best student paper award** JCDL 2008)
- Ding, H., Trajcevski, G., Scheuermann, P., **Wang, X.** and Keogh, E. *Querying and Mining of Time Series Data: Experimental Comparison of Representations and Distance Measures*. (VLDB 2008).
- Ye, L., **Wang, X.**, Yankov, D., and Keogh, E. *The Asymmetric Approximate Anytime Join: A New Primitive with Applications to Data Mining* (SDM 2008)
- Ye, L., **Wang, X.**, and Keogh, E. *Autocannibalistic and Anyspace Indexing Algorithms with Applications to Sensor Data Mining*. (SDM 2009)
- Kasetty, S., Stafford, C., Walker, P., **Wang, X.** and Keogh, E. *Real-Time Classification of Streaming Sensor Data*. (ICTAI 2008).
- Zhu, Q., **Wang, X.**, Keogh, E, Lee, S. and Rampley, T. *Towards Indexing and Data Mining all the Worlds Rock Art*. In the 37th Annual International Conference on Computer Applications and Quantitative Methods in Archeology (CAA 2009)
- **Wang, X.** and Keogh, E. *Finding Centuries-Old Hyperlinks with a Novel Semi-Supervised Learning Technique*. (JCDL 2009)
- Zhu, Q., **Wang, X.** and Keogh, E. *Augmenting the generalized hough transform to enable the mining of petroglyphs*. (KDD 2009)

## REFERENCES

References Available Upon Request.