



# UMass Lowell Computer Science Colloquium Announcement

**Speaker:** Dr. Zhixiang Chen  
Dept of Computer Science, UTPA

**Date:** March 11, 2002

**Time:** 3:00pm - 4:00pm

**Place:** Olsen Hall 311  
Refreshments are served at 2:30pm

---

## On the Capabilities of Outlier Formulation Schemes for Data Patterns with Diverse Characteristics

Outlier detection is concerned with discovering the exceptional behaviors of objects. Its theoretical principle and practical implementation lay a foundation for some important applications such as credit card fraud detection, discovering criminal behaviors in e-commerce, discovering computer intrusion, and etc.

In this talk, a unified model for several existing outlier detection schemes is first presented, and a compatibility theory is proposed to establish a framework for describing the capabilities for various outlier formulation schemes in terms of matching users' intuitions. Under this framework it is shown that the density-based scheme is more powerful than the distance-based scheme when a data set contains patterns with diverse characteristics. Density-based scheme, however, is less effective when the patterns are of comparable densities with outliers. A connectivity-based scheme is then proposed to improve the effectiveness of the density-based scheme when a pattern itself is of similar density as an outlier. The density-based and connectivity-based schemes are compared with each other and theoretical results are obtained in terms of their similarities, dualities, and strength and weakness. Empirical analysis is also given to demonstrate the applications with different features where each of the two schemes is more than effective than the other.

(This is the joint work with Jian Tang, Ada Wai-chee Fu and David W. Cheung.)

---