

# **Anomaly Detection via Over-sampling Principal Component Analysis**

**Yuh-Jye Lee, National Taiwan University of Science and Technology**

## **Abstract**

Outlier detection is an important issue in data mining and has been studied in different research areas. It can be used for detecting the small amount of deviated data. In this article, we use "Leave One Out" procedure to check each individual point the "with or without" effect on the variation of principal directions. Based on this idea, an over-sampling principal component analysis outlier detection method is proposed for emphasizing the influence of an abnormal instance (or an outlier). Except for identifying the suspicious outliers, we also design an on-line anomaly detection to detect the new arriving anomaly. In addition, we also study the quick updating of the principal directions for the effective computation and satisfying the on-line detecting demand. Numerical experiments show that our proposed method is effective in computation time and anomaly detection.