

Dr Shaoning (Paul) Pang
, KEDRI

spang@aut.ac.nz

Areas of supervision

1. Enterprise Data Warehousing, Data Mining and Knowledge Discovery Applications
2. Financial and Business Intelligence
3. Speech & Image Processing, Computer Vision, and Pattern Recognition
4. Mobile Robotics
5. CAD data analysis for Art and Design
6. Machine Learning for Internet Security

1. Research Topics on String Data Analysis

String Discrimination and Text Classification with Application to Internet Security: Latent Dirichlet Allocation (LDA) is a Bayesian model in the spirit of string discrimination that models each data point (e.g., a document) as a collection of draws from a mixture model in which each mixture component is known as a topic. The proposed research is to conduct LDA analysis over different types of network string data and better decision making for internet security. The student is required to conduct literature review, code LDA algorithms using Matlab or other programming language, and experiment LDA with internet security problems, such as spam or intrusion detection and classification.

String Kernel and String Classification for Network Traffic and Memory Analysis : Existing methods for attack detection requires an attack signature determined in advance. For unknown attack detection, the proposed research is to develop a network traffic and memory analyser capable of detecting an unknown attack without earlier determined attack signatures. The proposed method is a string kernel based SVM (support vector machine) supervised machine learning. The student is required to code several types of string kernel and enable the constructed kernels to work with SVM (software is provided) for string data classification.

Ontology Learning and Visualization above Google Information Retrieval : Ontology learning is a promising technique for automated knowledge acquisition from text corpora. The proposed study is to construct an ontology representation for Google search, reflecting the similarity between the query key words and the search results in 3D space. The student is required to conduct literature review, code ontology learning algorithms using Matlab or other programming language, and experiment his algorithms with synthetic string data, text document data, as well as data from Google search engine.

For each of the above 3 research topics, a maximum 5000NZD scholarship is provided for a 90points Master thesis study under the FY2009 research program of KEDRI-NICT project. (Conditions applied, please contact Dr. Pang for the detail.)

2. Research Topic on Image and CAD Data Analysis (associated with KEDRI-NICT project)

A technical interpolation of art information from images: This research is about developing the techniques of translating the `art' information that is contained/embedded in our real life images, into a set of art components, such as a piece of contour/curve description, a type of texture and colour pattern, or a 3D structure that are the most valuable for the art and design. Most importantly, they will be able to be used by the art designer in future design work. The student will be required to construct an image database, develop two or three types of art information interpolation methods, and apply the interpolated components from the image database to a new art work design.

	<p>3. Research Topic on Financial Data Analysis</p>
--	--

Geometrical Correlation Analysis for Long Term Market Trend Prediction: Upon financial data including stock market prices, commodity prices, time bargain prices, and currency exchange rates, has increasingly affected our daily life, people are more interested than ever to know the trend of the market. The proposed research is on geometrical correlation analysis and correlation-aided long term market trend prediction. The student is required to develop new correlation analysis and correlation-aided prediction methods, and experiment developed methods with provided various foreign currency exchange historical data.