

Seminar: Friday 28 March 2014:

Monash Undergraduate Research Projects Abroad (MURPA)

Student Presentations live – Including High Definition Video links with the United States, Singapore and the United Kingdom

The MURPA program supports a unique summer research placement for undergraduate students studying computer science, software engineering, or technically oriented IT units, in a leading research group overseas. It not only provides a research experience at the undergraduate level, but does that in an international context. Students undertake an 8-week summer semester international research project at the prestigious University of California, San Diego (UCSD) and the University of Warwick, UK.

This annual program exposes final year undergraduate students to an international research experience within a leading research laboratory.

This year for the first time saw participation in this program by the University of Queensland (UQ), with one student being placed with a host at the UCSD. And continuing with our theme of firsts, this year saw two students from Monash placed at the Institute for Infocomm Research (I2R) in Singapore.

On Friday 28 March the group of returned students will present their outcomes by High Definition Video back to their international host and live audiences at Monash Clayton campus, and to the UQ St Lucia campus.

Overview of Projects:

James MacIndoe (Bachelor of Science and Bachelor of Engineering)

Project title: WiFire Android App

Abstract: WiFire is a research project at UCSD aimed at the real-time simulation, predication and visualisation of wildfires. For my project, I developed an Android app for WiFire. The app provides data visualisation using an overlay on Google maps and can issue alerts to the user when certain user-specified conditions are met, such as weather conditions.

Brian Song (Bachelor of Engineering (Software), MPhil candidate, (University of Queensland)

Project Title: Temporal Dynamics of Learning Conference (TDLC) and rat-robot Experiments at UCSD

Abstract: Since US President Obama announced the US government funded Brain Initiative in early 2013, there has been a consolidated and formal recognition of the importance and far-reaching impacts of advancing brain research. This brief seminar will provide an overview of the TDLC conference, including my own poster presentation, as well as the rat-robot experiments that were conducted at the UCSD Cognitive Science Department.

Jie Liu (Bachelor of Software Engineering)

Project Title: CIPRES Gateway Usage Analysis

Abstract: The CIPRES Science Gateway hosts thousand users each year. These users submit hundreds of thousands jobs to XSEDE resources each year. Both job record information and metadata about each job run are stored in the CIPRES MySQL database. The database contains information about all submitted jobs that have been run over the course of the 4 years of operation. Currently the job data analysis is mainly done by manually manipulating a .csv file, which contains both the job record information and

metadata. So there is a need to reduce the pain from the current tedious data analysis process. The project goal is to create a software package which allows users who may not know SQL to have access to CIPRES database via a provided GUI with the software package. More specifically, the GUI allows users to select the use cases, which is based on the predefined questions of the job data by users. For each use case, there are multiple queries associated with it. Every time, users just need to fill the fields related to each use case, for example, the time range, the user email address etc., then after users confirm these filled information, the software package will run those associated queries and finally return results as tables. Moreover, the software package could also generate tables with details of each job. As required, the detailed information contained in tables could also be exported into a .csv file for further processing.

Jonathan Samosir (Bachelor of Computer Science)

Project Title: A real time data stream processing system for enabling patient tailored healthcare

Abstract: Research is currently being done on a new platform to deliver healthcare tailored to the individual: SenseHealth. For this platform to exist, both traditional electronic medical records (EMR) and medical data being streamed in real-time from patient sensors need to be processed to develop models that will go on to formulate a given patient's healthcare plan. My time was spent evaluating, working with and modifying systems that will go on to be the part of SenseHealth that will be receiving and processing medical data streams coming from patient health sensors. The most significant of which was the Apache Samza distributed stream processing framework, developed at LinkedIn. The short talk will go into details of my experiences over the 12 weeks working with Samza, both as a user and contributor, trying to get it to fit into the model of the needed data stream processing framework, working with the San Diego Supercomputer Center and FutureGrid cyberinfrastructure, and my MURPA experience as a whole. I will also touch on future research possibilities that have been made available through being involved in this MURPA project.

Roy Hyde (Bachelor of Software Engineering)

Project Title: Adaptive Classification for High-Speed Data Streams

Abstract: Many modern applications generate large amounts of data at very high speeds, which cannot be analysed by traditional data analytics techniques in real time. The goal of data stream analytics is to extract hidden knowledge and patterns from data streams, whilst addressing the challenges inherent with data streams. Some of these challenges include only being able to perform a single-pass over the data, the need to deliver results in real time, and detecting the occurrence of drift within the stream. This objective of this project is to learn more about the decision tree classifier family of data stream algorithms in order to make improvements to the speed and accuracy at which they perform.

Travis Liew (Bachelor of Software Engineering)

Project Title: Pairs Trading vs. Clusters Trading

Abstract: The research conducted at I2R Singapore involved investigating the merits of investing based on clusters of stocks versus the market neutral strategy Pairs Trading, which trades on pairs of stocks. To achieve the clustering of stocks, two techniques were explored; Affinity Propagation and Spectral Clustering. In doing so, we hoped to achieve an improvement on returns to that of Pairs Trading. We attempted to closely replicate both the methodology and the Pairs Trading strategy employed in the experiments of two well-known research papers to ensure conclusive results when comparing to that of our clustered trading strategy, deemed Clusters Trading.

Ahmed Shifaz (Bachelor of Software Engineering)

Project Title: Optimizing Algorithm for Inferring Causal Structure of Gene Regulatory Network (GRN)

Abstract: The aim of this project is to optimize an algorithm written by Christopher Penfold at University of

Warwick, doing bioinformatics research. His paper "Non-parametric Bayesian inference for perturbed and orthologous gene regulatory networks" (C.Penfold, D.Wild, et al, 2012) describes an algorithm to infer GRN's based on time-series data. He currently has a matlab implementation of the described algorithm. However it requires significant amount of time to run larger gene networks using the current implementation. Therefore the project aims to optimize this matlab implementation by porting part of it to "C", finding methods to reduce the time required to infer GRN's using other approximation techniques, and to explore accuracy of such approximations.

If you would like to learn about the exciting results they can achieve, learn about their experiences and more about the program then please join us on Friday.