From:	nazrina@comp.nus.edu.sg
Sent:	Wednesday, August 30, 2006 12:30 PM
To:	seminar-l@comp.nus.edu.sg
Cc:	lttech@comp.nus.edu.sg; nazrina@comp.nus.edu.sg
Subject:	CS SEMINAR by Associate Professor Srinivasan Parthasarathy Srinivasan

Parthasarathy
MIME-VERSION: 1.0
Content-type: text/plain

(http://www.comp.nus.edu.sg/cs/csseminar.html)

NATIONAL UNIVERSITY OF SINGAPORE

School of Computing

C S   S E M I N A R

TITLE : Mining Protein-Protein Interaction Networks

SPEAKER : Associate Professor Srinivasan Parthasarathy Srinivasan
          Parthasarathy
          Computer Science and Engineering Department
          Ohio State University (OSU)

TIME : August 31, 2006, 3:00pm - 4:00pm, Thu

VENUE : EC - Executive Classroom, SoC1, Level 5 Room 46
        School of Computing, National University of Singapore

Chaired by Dr Tung Kum Hoe, Anthony (atung@comp.nus.edu.sg)

ABSTRACT:

In this talk I will describe our recent work on analyzing protein protein interaction
networks. The objective is to find clusters of proteins that perhaps have common
functionality. Such clusters can then be used to potentially identify novel functions of
un-annotated proteins. From an algorithmic perspective a principal challenge is that the
graphs embedding such interactions are scale-free. This makes it less than amenable to
standard clustering or graph partitioning approaches. A further complication is that it is
also believed that the current state of knowledge about such graphs is incomplete in the
sense that many of the interactions currently reported in the literature are believed to
be false.

Again clustering methods that are sensitive to noise and outliers do not work well in this
context. In this talk I will describe specific solutions that we have developed for this
problem, particularly as they apply to the protein-protein interaction network of yeast.
These
include: i) novel yet simple preprocessing steps to identify, detect and eliminate
potential false positives; ii) hub duplication, to alleviate the impact of hub nodes on
graph clustering or partitioning and to enable soft clustering of proteins and their
interactions based on the notion of dense proximity sub-graphs, and iii) ensemble
clustering using different topological measures to improve the robustness of the algorithm
to noise.

The strategies evaluated significantly improve the quality of the resulting clustering as
evaluated by a range of statistical, information theoretic and domain specific quality

measures. This is joint work with my graduate students Sitaram Asur and Duygu Ucar.

BIODATA:

Srinivasan Parthasarathy Srinivasan Parthasarathy is currently an Associate professor at the Computer Science and Engineering Department at the Ohio State University (OSU). He heads up the data mining research laboratory and has a joint appointment in the Department of Biomedical Informatics at Ohio State. He received a B.E in Electrical Engineering from the University of Roorkee (now IIT-Roorkee) and an MS and PhD in Computer Science from the University of Rochester. His research interests include data mining, high performance computing & systems, scientific data analysis and bioinformatics. He is a recipient of the US National Science Foundation's CAREER award, the US Department of Energy's Early Career Principal Investigator Award, and an SBC/Ameritech Faculty fellowship. His work has received several awards including an IEEE Data Mining 2002 best paper, a SIAM Data Mining 2003 best paper, the VLDB 2005 best research paper and a "Best of SDM05" selection from SIAM Data Mining 2005.