Volume 3, Number 5 June 2016

Interdisciplinary collaboration between engineering, mathematics and science

SEMS Research Highlights

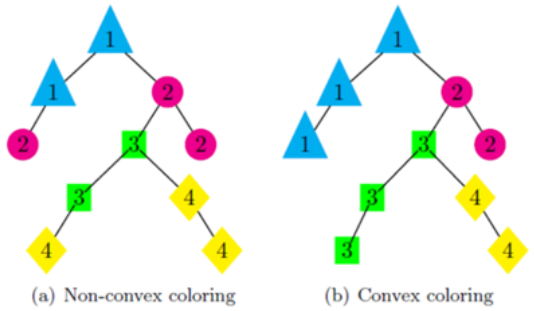
Linear Programming Approach to Clustering a Biological Network into Connected Subnetworks

Dr. Sangho Shim

Assistant Professor of Industrial and Manufacturing Engineering, Engineering Department

This newsletter presents the research conducted within the School of Engineering, Mathematics and Science (SEMS) at Robert Morris University (RMU). It covers various relevant topics including: interdisciplinary efforts, successful research grants, student research, posters and papers, journal publications, presentations at national and international conferences, contribution to professional societies, STEM educational research, industrial consulting collaborations and applied research.

Clustering a large scale network into connected subnetworks is an important problem in bio-informatics, which is defined as computational technology to analyze large scale biological data to simplify and visualize a large scale biological network such as a phylogenetic tree and a protein-protein interaction network. In the last decade, Israeli computer scientists introduced an optimization problem, the so-called convex recoloring problem to encode the clustering problem and have developed straightforward algorithms such as dynamic programming algorithms and approximation algorithms that are specifically designed for solving this problem. The convex recoloring problem is to change the initial colors at the minimum number of nodes of a phylogenetic tree such that the induced subgraph of the nodes colored with each color is connected. The convex recoloring problem is the underlying problem of various clustering problems raised not only in phylogenetics but also in transportation science, computer networks, linguistics and many other fields.



(a) Non-convex coloring (b) Convex coloring

The subnetwork connectivity is depicted in the adjoining coloring diagram where nodes of the same color are shown connected in (b) using the improved algorithm.

Linear programming (LP) is a computational scheme to model and solve an optimization problem as a linear programming formulation defined by a system of linear inequalities along with a linear objective function to optimize over the solutions to the system. The LP formulation of the convex recoloring problem is an integer linear programming (ILP) formulation indicating each clustering by an integer point in the polyhedron defined by the system of linear inequalities. Current work improves the conventional model of the convex recoloring problem using an extended formulation. It is shown mathematically and computationally that the proposed innovative LP model can be solved much faster than the conventional LP model as well as the previously known straightforward algorithm. A phylogenetic tree of thousands nodes can thus be efficiently clustered into connected subtrees by LP.

This is a publication of SEMS - Research and Outreach Center (ROC) which was established in 2010 by the SEMS Dean Dr. Maria Kalevitch. SEMS-ROC connects SEMS faculty and students with the region, the nation and the globe, demonstrates diversity and interdisciplinary interests of all three departments in the school. For more information on research at RMU – SEMS please contact:
Dr. Priyadarshan Manohar,
Co-Director, SEMS-ROC, Research and Grants, E-mail: manohar@rmu.edu, Tel.: 412 397 4027