

## **PageRank-flux On Graphlet-Guided-Network(PRO-GGNet): A Method for Pathway Reconstruction and Multi-Omic Data Integration**

Kaan Arıcı m.kaanarici@gmail.com Turkey METU

Nurcan Tunçbağ ntuncbag@ku.edu.tr Turkey Koc University

The recent advancement of omic technologies provides snapshots of cells, tissues, or patients identifying prominent genes, proteins, metabolites, and small molecules. However, accumulated big data on various omic data types may inherently make diseases or perturbations incomprehensible. Network inference or reconstruction methods map a set of significantly altered proteins/genes/metabolites to a given reference network that is composed of already known relations or interactions. Followingly, the signals from these significant hits are propagated so as to identify biomarkers, drivers, pathways or disease modules consisting of a relatively small number of genes. Including Steiner trees/forest, random walk, or heat diffusion, network inference or reconstruction approaches are mainly based on global and local topological features of networks considering graph theory. However, the problem of identifying interacting or related genes in an extensive reference network and assessing associations between multiple sparse omics data is one of the critical problems in system biology. A network motif is a significantly frequent subgraph pattern in a given network. In this study, using graphlets (small non-isomorphic subgraphs) as network motifs, we reduce the dimension of the reference network dramatically and focus on the functionally important regions of the network. However, identifying network motifs and saving their knowledge in an extensive network requires high computational costs, limiting the usage of graphlet motifs for real problems. Our approach, PageRank-flux On Graphlet-Guided-Network (PRO-GGNet), focuses on the motifs of a given gene-set in an extensive network and constructs the graphlet-guided network to reduce the computational cost followed by applying PageRank algorithm. Next we compared the performance of PRO-GGNet, with OmicsIntegrator and PathLinker that use prize collecting Steiner forest solution and shortest path extraction methods, respectively. PRO-GGNet outperformed the tools in testing the inference of Immune and Cancer Signaling Pathways in NetPath by providing the highest precision, recall and F1-scores.