



Jason Liu, Min Xu, Jinchuan Xing

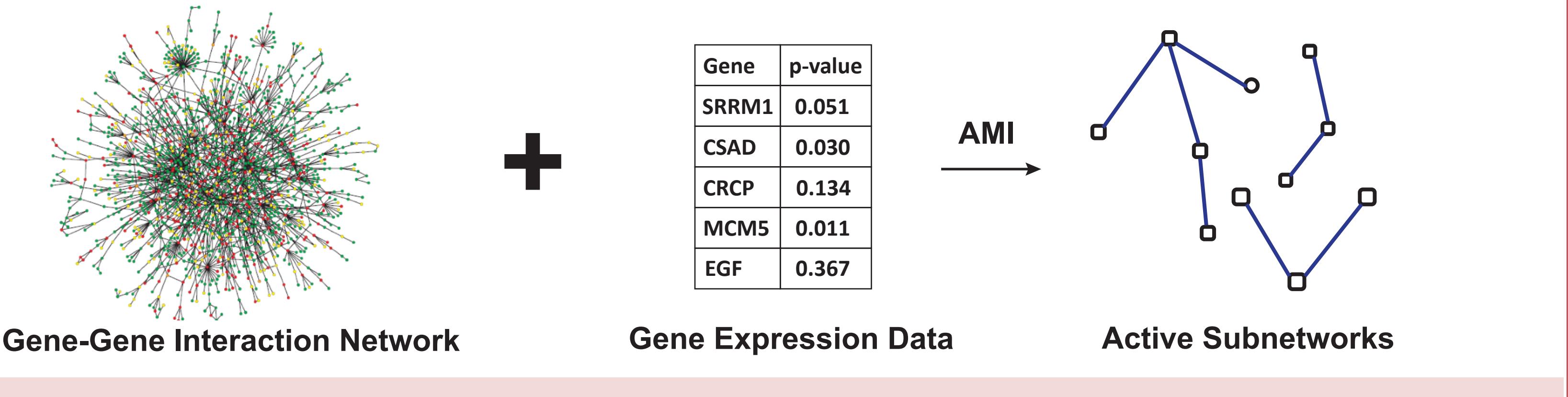
Rutgers, the State University of New Jersey, Piscataway, NJ, USA

## Introduction

- Current network analysis relies on filtering and literature
- Active Module Identification (AMI) can be used to identify important parts of the network representing biological pathways
- Many AMI algorithms have been proposed but the selection of their use can be arbitrary
- We evaluate four algorithms across four datasets to better understand their use-cases and advantages

## Methods

### Active Module Identification (AMI)

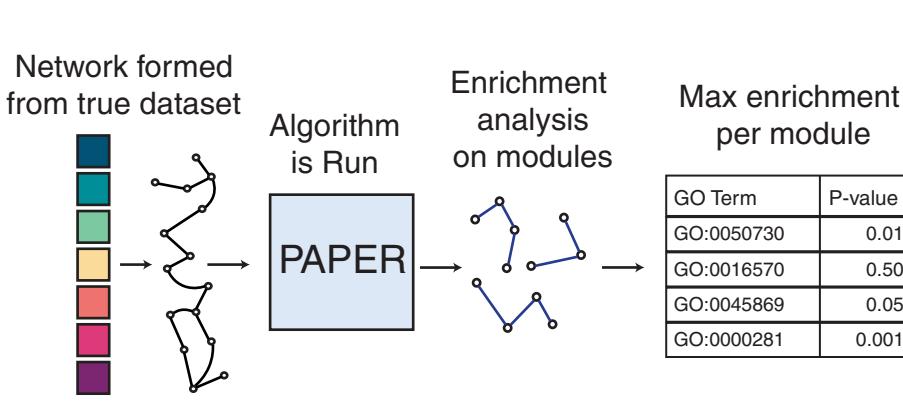


### Study Design

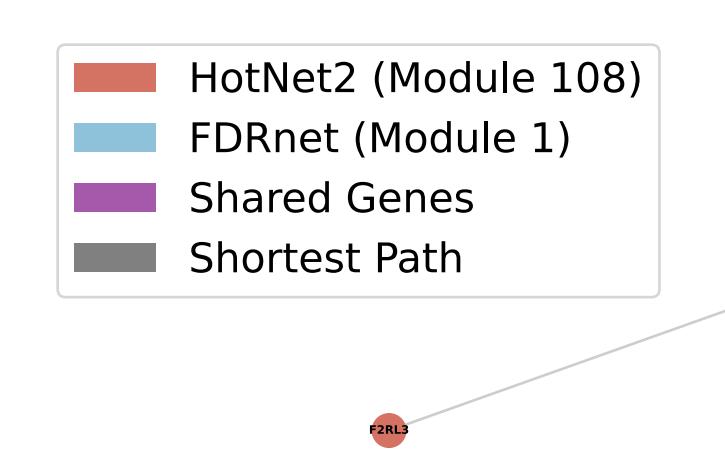
- Four algorithms:
- PAPER [1]
  - DOMINO [2]
  - HotNet2 [3]
  - FDRnet [4]

- Four datasets:
- Illumina [5]
  - IonTorrent [6]
  - TnFA [7]
  - Fly Transcriptome [8]

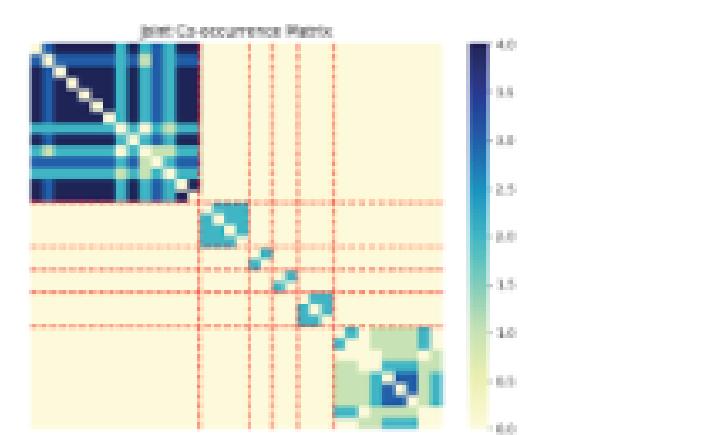
#### 1. Evaluation with EMP [2]



#### 2. Earth Mover's Distance for similarity analysis

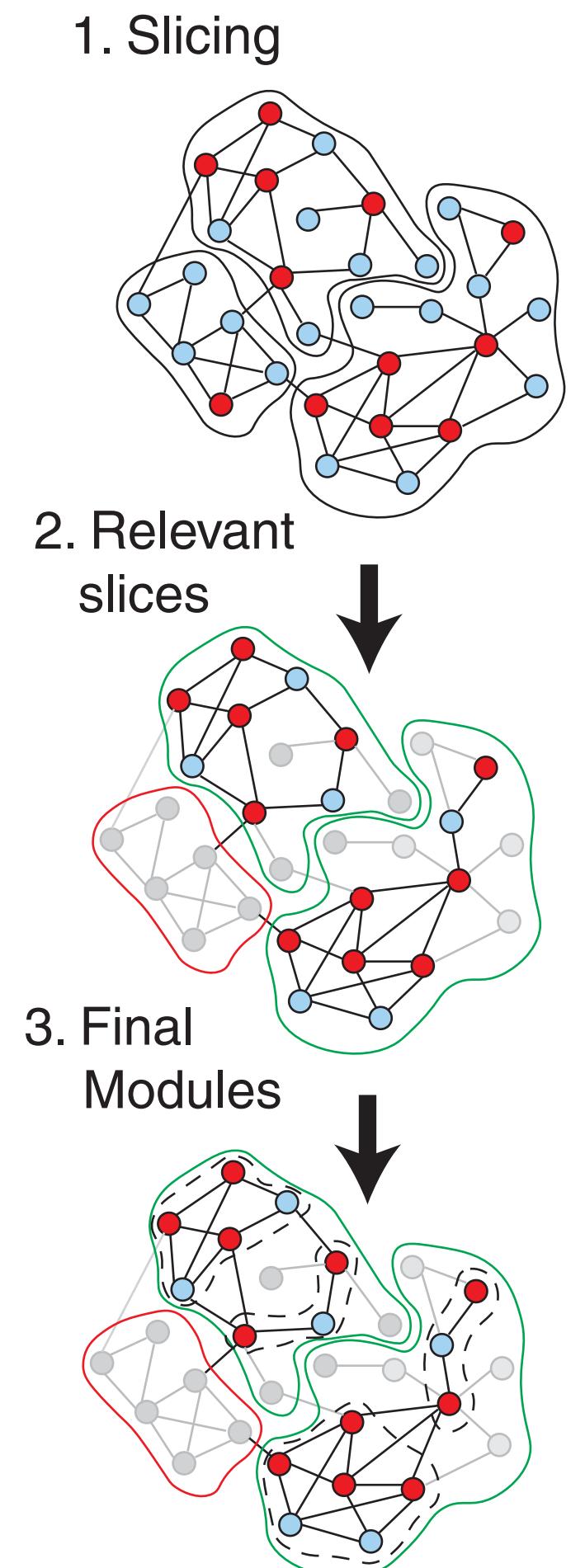


#### 3. Spectral Clustering on co-occurrence matrix

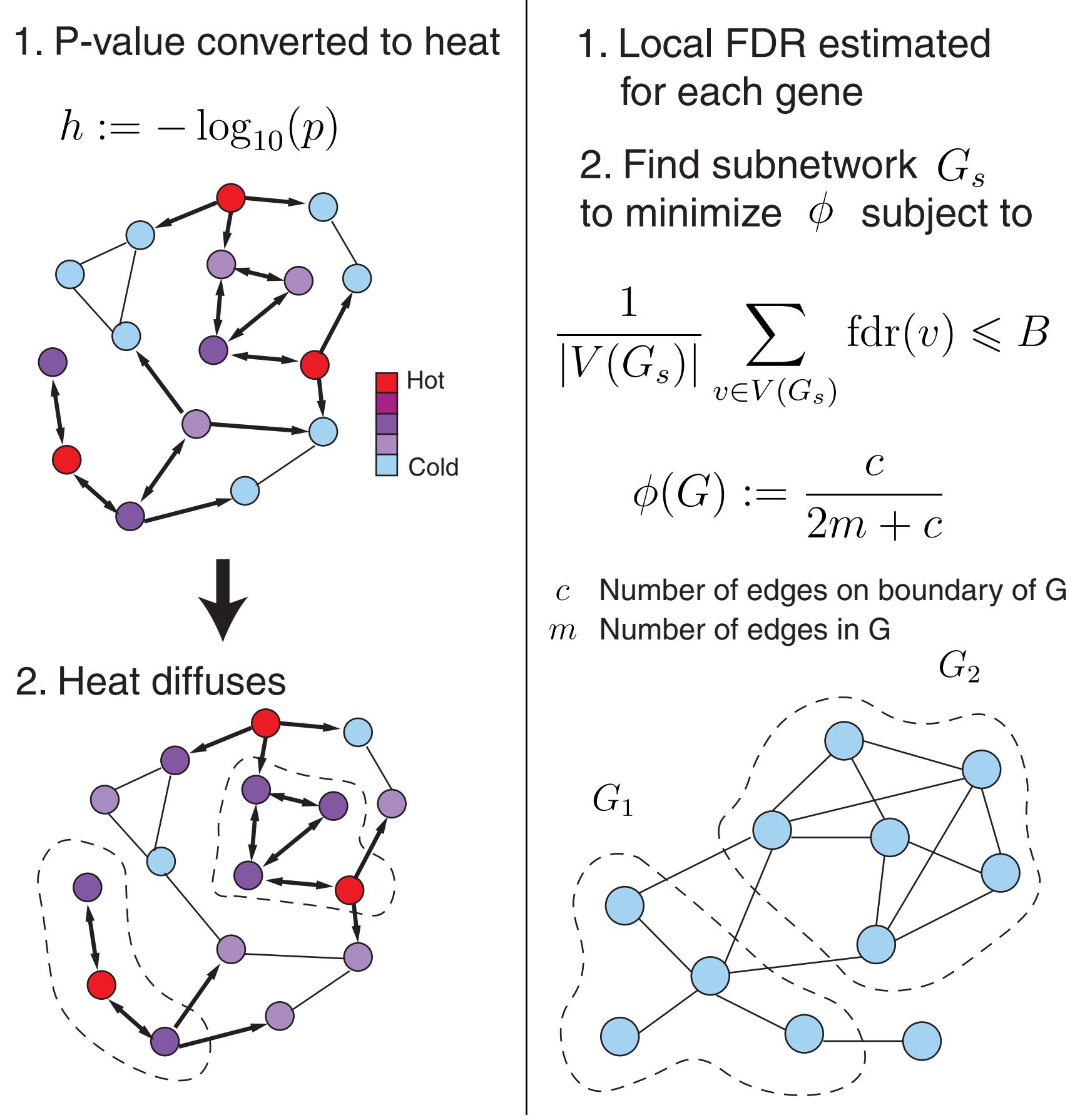


### Algorithms

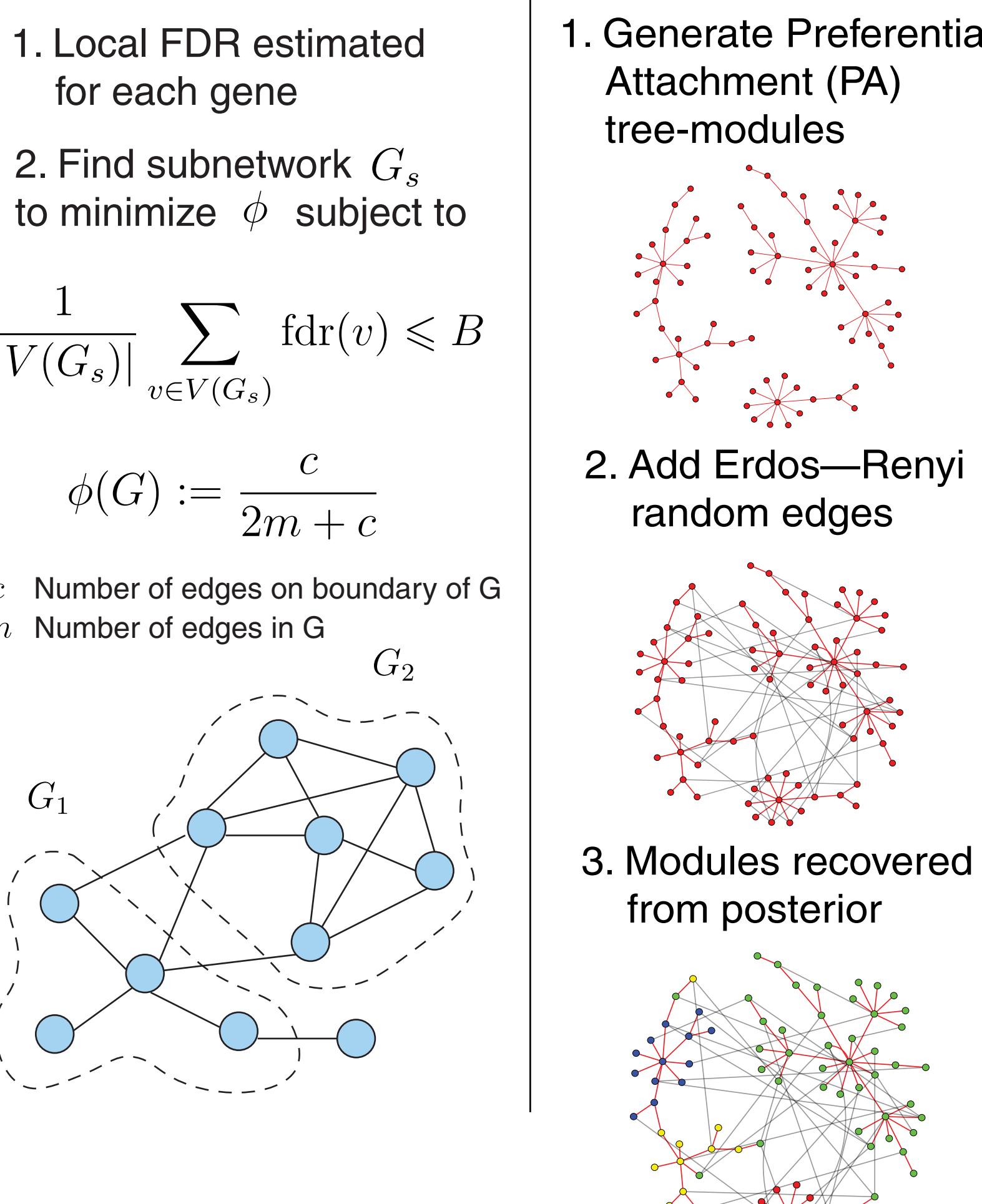
#### DOMINO



#### HotNet2



#### FDRnet



#### PAPER

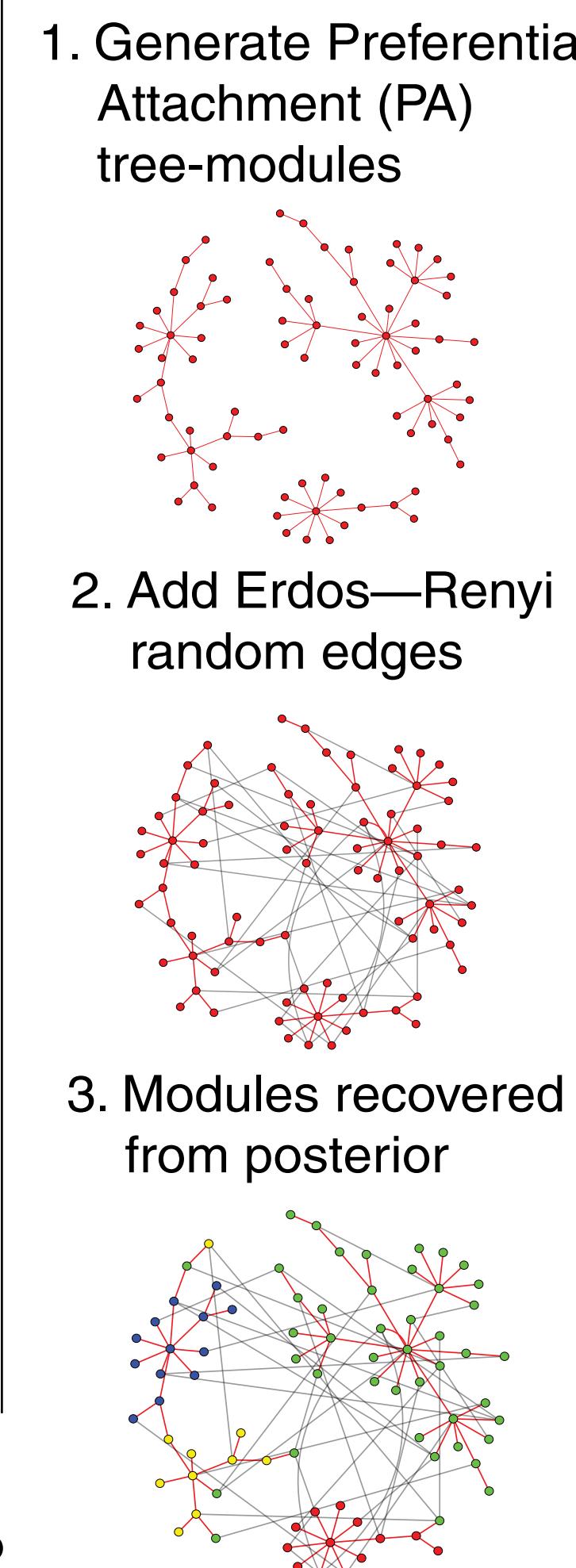


Figure 1. DOMINO, HotNet2, FDRnet, PAPER

## Performances of Algorithms

Algorithm performance depends on dataset, no overall “winner”

Algorithms can perform well, even if they produce very different sizes or counts of modules

	Illumina	IonTorrent	TnFA/DIP	Fly transcriptome
PAPER	<b>0.938 (15/16)</b>	0.83 (5/6)	<b>1 (48/48)</b>	<b>1 (275/275)</b>
DOMINO	0.606 (125/206)	0.702 (149/202)	0.987 (239/242)	0.864 (377/436)
HotNet2	0.004 (4/903)	0.01 (15/1501)	0.976 (163/167)	0.065 (60/923)
FDRnet	0.9 (18/20)	<b>1 (8/8)</b>	<b>1 (33/33)</b>	0.533 (24/45)

Table 1: EHR and enrichment per algorithm

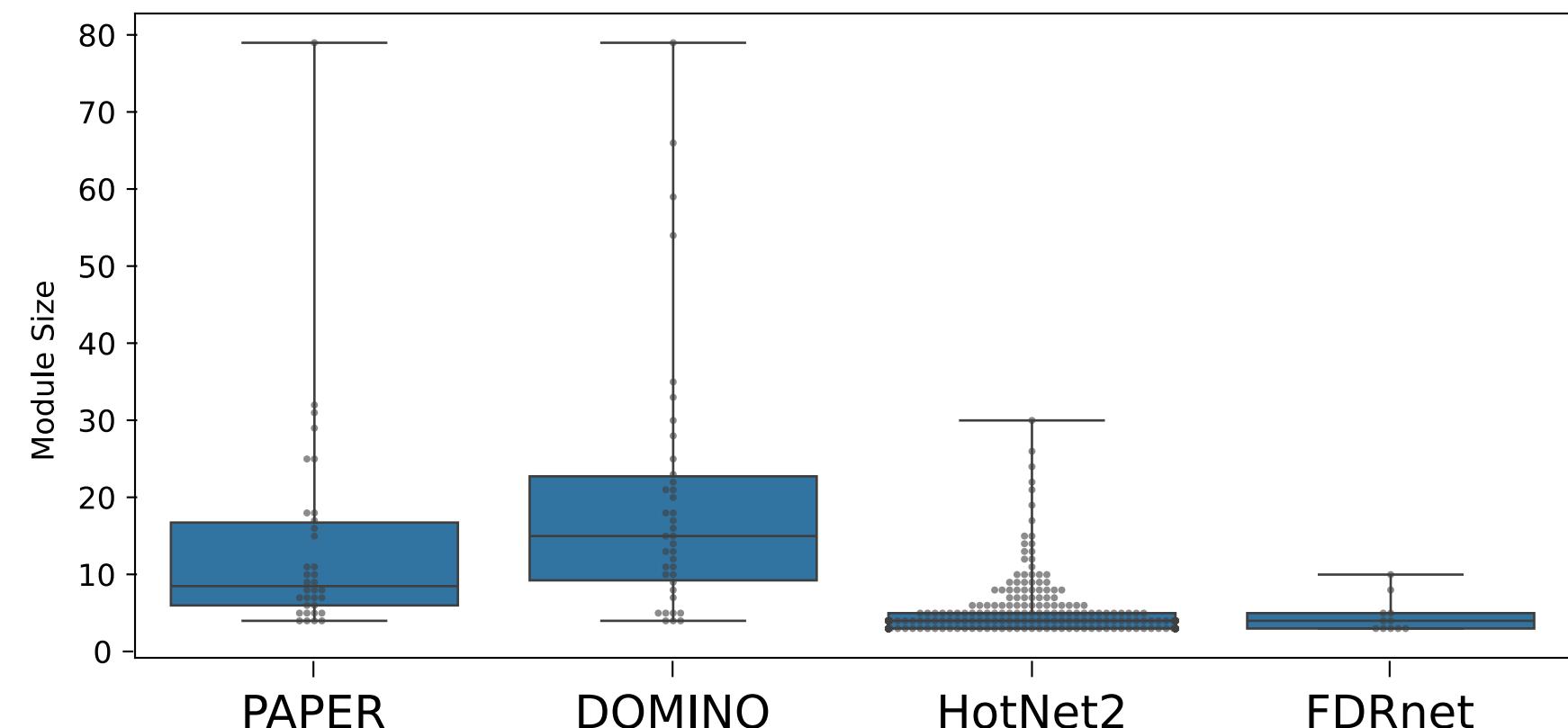


Figure 2: Module size distribution per algorithm

## Module Similarity Analysis

Motivating Question: Many algorithms perform well, but are they similar?

We use the *Earth Mover's Distance* to measure the similarity of non-overlapping modules

