Bipartite Cohesive Subgraph Discovery in Networks

Minsik Lee | Woocheol Shin | Kyuhwan Lee | Sunwoo Yeom | Jaeyun Shim | Seongsik Hwang

Prof. Junghun Kim

ABSTRACT

In an era dominated by extensive internet usage and the proliferation of social networking platforms, understanding the intricacies of social networks becomes paramount. This paper underscores the significance of identifying cohesive subgraphs, particularly in bipartite networks—comprising two distinct node sets. Through a comprehensive assessment of various cohesive subgraph models on both real-world and synthetic datasets, this study makes two pivotal contributions: consolidating and implementing existing models into a single accessible framework, and offering an in-depth analysis across diverse networks, emphasizing bipartite structures. This exploration is poised to benefit applications ranging from recommender systems to collaboration networks.

INTRODUCTION

This study focus on bipartite graphs, which are simple, non-weighted, and undirected.

Motivation: the proliferation of internet services and mobiles has boosted social networks. As interest in analyzing these networks grows, identifying cohesive subgraphs becomes vital for community detection and recommendations.

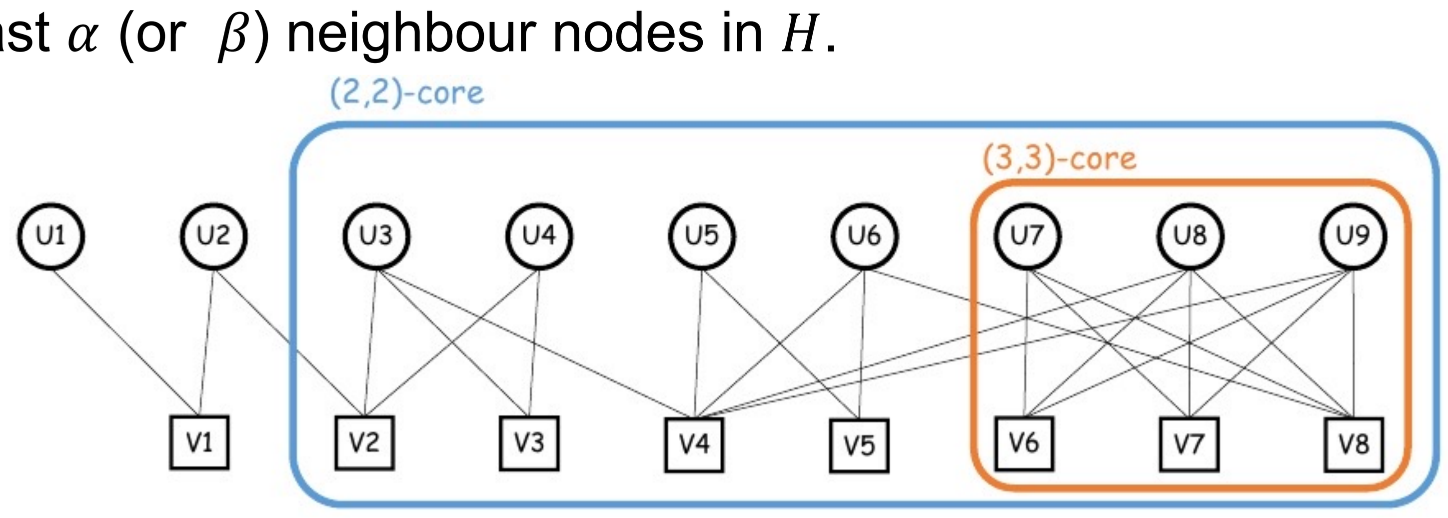
Challenges: The two key challenges are as follows.

- Consolidation and Implementation of Existing Models: We have taken the initiative to implement various cohesive subgraph models and amalgamate them into a singular framework.
- In-depth Analysis on Varied Datasets: We present comprehensive experimental findings using a wide array of networks, with particular emphasis on bipartite networks. The analysis aims to pinpoint the most suitable cohesive subgraph model for different situations.

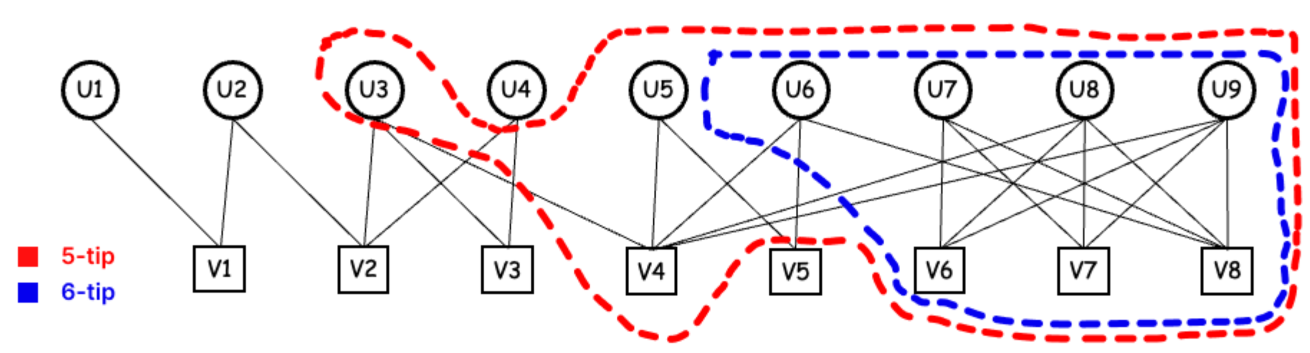
BIPARTITE COHESIVE SUBGRAPH MODELS

These are definitions of the bipartite cohesive subgraph models used in this paper.

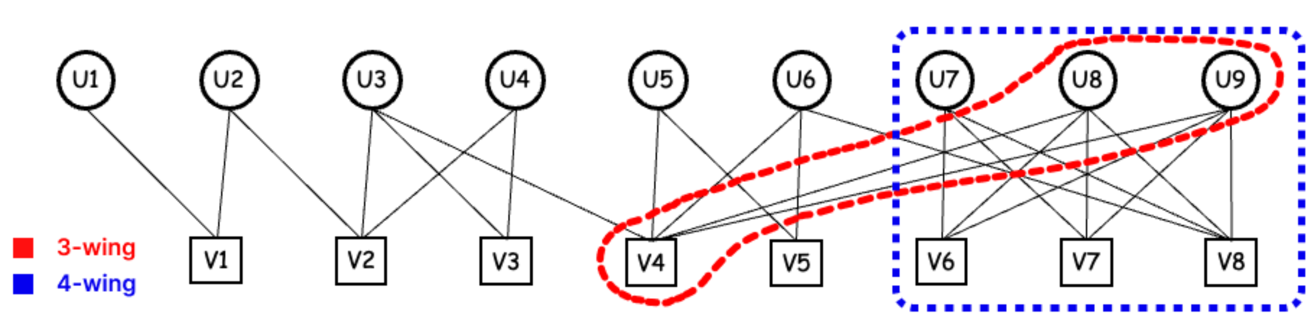
(α,β)-core: Given a bipartite graph $G = (U, V, E)$ and positive integer α, β, $(\alpha,\beta) - core$ is a set of nodes H of which every node H_U (or H_V) has at least α (or β) neighbour nodes in H.



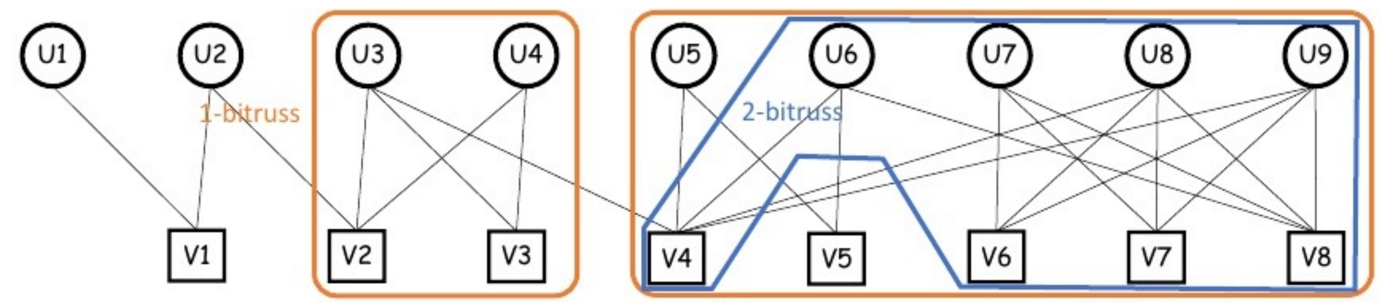
k-tip: Given a bipartite graph $G = (U, V, E)$ and positive integer k, a subgraph A bipartite subgraph $H = (U, V, E)$ of G, induced on U, is a $k - tip\ iff$ each vertex $u \in U$ takes part in at least k butterflies, each vertex pair $(u, v) \in U$ is connected by series of butterflies and H is maximal, i.e., there is no other $k - tip$ that subsumes H.



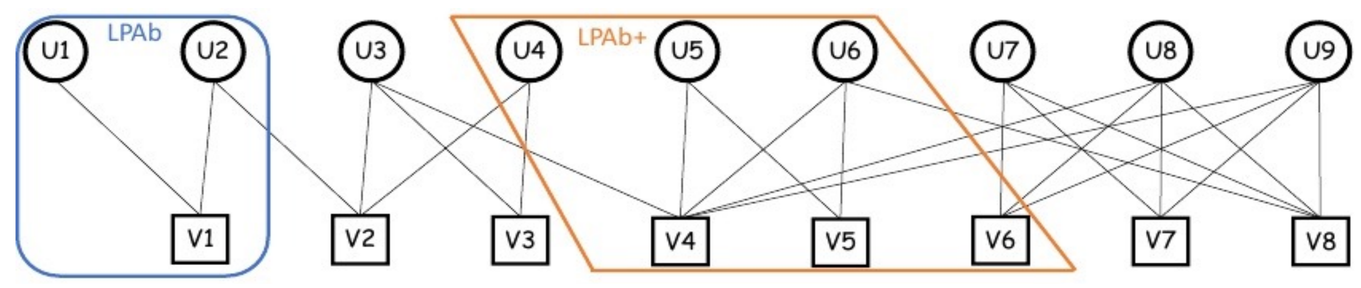
k-wing: A bipartite subgraph $H = (U, V, E) \subseteq G$ is a $k - wing$ if each edge $(u, v) \in E$ takes part in at least k butterflies, each edge pair (u_1, v_1), $(u_2, v_2) \in E$ is connected by series of butterflies and H is maximal, i.e., there is no other $k - wing$ that subsumes H.



k-bitruss: Given a bipartite graph $G = (U, V, E)$ and an intenger $k \geq 0$, the k-bitruss of G, denoted by $TK\ (G)$, is the largest edge-induced subgranph H of G such that every edge of H is contained in at least k rectangles within H.



LPAb: The LPAb method for bipartite graphs assigns unique labels to nodes, which are updated to enhance modularity. The iterative process stops when no further increase in modularity is possible, and nodes with the same label are recognized as a community.



RESULTS

Dataset: We utilise synthetic networks to evaluate various models.

Synthetic network configuration:

BNOC1 to BNOC3: μ = 0.6, n = 0.1, x = [4, 4], z = [3, 3], $p = null$.

BNOC4 to BNOC7: μ = 0.3, n = 0.1, x = [4, 4], z = [3, 3], $p = null$.

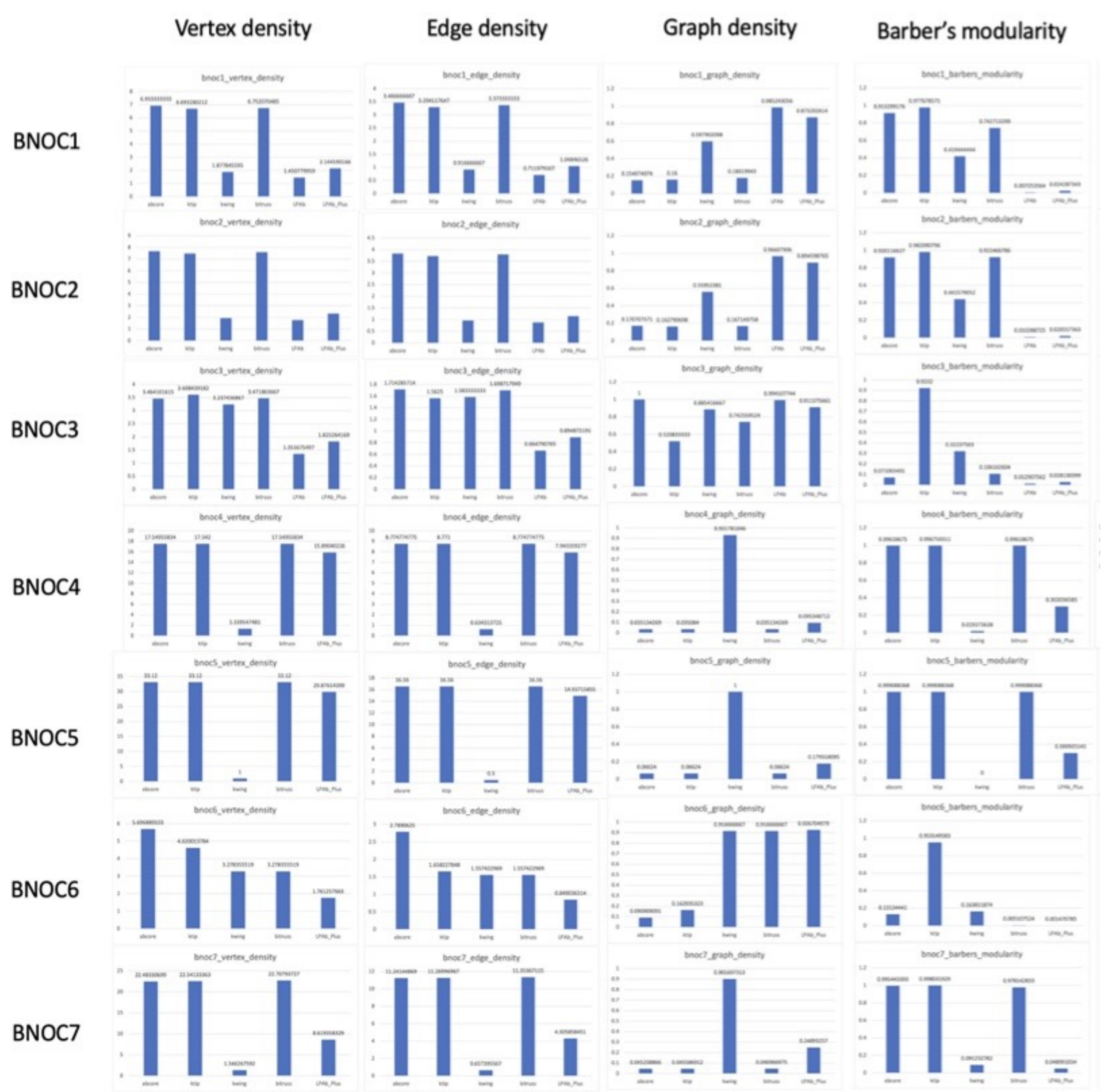
Evaluation Metrics:

Vertex Density: $|E|$ / $\sqrt{(|U| \times |V|)}$.

Edge Density: $|E|$ / $(|U| + |V|)$.

Graph Density: $|E|$ / $(|U| \times |V|)$.

Barber's Modularity: $(|E|G - |E|H)$ - $(|U|H \times |V|H)$ / m^2.



CONCLUSION

Navigating Social Networks: Insights into Bipartite Networks

Our study delves deep into social networks, revealing key insights into bipartite networks and cohesive subgraph detection. By merging various subgraph models into one toolset, we've simplified future research in this area. Our extensive analysis highlights the importance of understanding these networks, opening avenues for improved applications across various fields.