A Novel Higher-Order Framework for Improving Community Detection in Social Networks*

Rodrigo Hernández[1], Inmaculada Gutiérrez[1,2], and Javier Castro[1,2]

[1] Universidad Complutense de Madrid, Madrid, Spain
[2] Instituto Universitario de Estadística y Ciencia de Datos, UCM, Madrid, Spain

Abstract. This contribution highlights the work published in the *International Journal of Computational Intelligence Systems* (2025) [3]. We present the Multistructural Communication Graph Community Detection Problem, a framework that improves any matrix-based community detection algorithm by constructing a composite adjacency matrix that integrates direct, two-hop, and three-hop path information with optimised weights, found by grid search over more than five thousand combinations. On a benchmark of 150 real networks, our approach improves modularity in the majority of experiments, confirming that higher-order graph structure systematically enhances community detection quality across diverse network types and base algorithms.

Keywords: Social network analysis · Community detection · Higher-order matrices · MCG-CDP · Modularity

1 Introduction

Community detection is a fundamental task in social network analysis (SNA): given a graph $G = (V, E)$, the goal is to partition the node set V into groups that are more densely internally connected than to the rest of the network [2]. Communities reveal the latent functional organisation of complex systems across diverse domains.

The most widely adopted community detection algorithms, including the Louvain algorithm [1] and its refinement Leiden [6], operate on the adjacency matrix A of the undirected graph and maximise Newman–Girvan modularity [5]. Despite their efficiency, both methods are susceptible to resolution limitations and greedy local optima [4]: in sparse or hierarchically organised networks, the first-order adjacency matrix alone fails to encode the full connectivity landscape, and optimisation over A frequently misses structurally deeper patterns.

The work summarised here addresses this limitation by proposing the Multistructural Communication Graph (MCG) framework [3]. Given the adjacency matrix A of the network, the framework enriches the algorithm's input with two-hop and three-hop path information encoded in A^2 and A^3, where A^k counts the

* Project PID2024-155289NB-I00 funded by MICIU/AEI/10.13039/501100011033 and ERDF/EU.

number of walks of length k between each pair of nodes. The obtained composite matrix A^* captures both local (one-hop) and non-local (multi-hop) connectivity in a single, tunable representation. Crucially, the method is algorithm-agnostic: it acts as a transparent plug-in applicable to any matrix-based community detection approach.

2 The MCG-CDP Framework

2.1 Modularity and the Community Detection Problem

Let $G = (V, E)$ be an undirected graph with n nodes, m edges, and adjacency matrix $A \in \{0,1\}^{n\times n}$, where $a_{ij} = 1$ if $\{i, j\} \in E$. The matrix power A^k counts the walks of length k between every pair of nodes.

The Newman–Girvan modularity of a partition $\mathcal{P}$ of V is [5]

$$Q(A, \mathcal{P}) = \frac{1}{2m} \sum_{ij} \left[a_{ij} - \frac{k_i k_j}{2m} \right] \delta(c_i, c_j), \tag{1}$$

where k_i is the degree of node i, c_i the community of node i, and δ the Kronecker delta. High modularity ($Q \gtrsim 0.3$) indicates a partition with substantially more internal edges than expected under a random null model.

2.2 Composite Matrix and Optimisation

The MCG framework enriches the input seen by a community detection algorithm by combining three structural views of the network: direct adjacency (A), two-hop reachability (A^2), and three-hop reachability (A^3). The MCG-CDP framework constructs the parametric composite matrix

$$A^* = \alpha_1 A + \alpha_2 A^2 + (1 - \alpha_1 - \alpha_2)\, A^3, \tag{2}$$

where $\alpha_1, \alpha_2 \in [0, 1]$ with $\alpha_1 + \alpha_2 \leq 1$. Given a base community detection algorithm ALG, the MCG-CDP solution is

$$Q^* = \max_{\alpha_1,\, \alpha_2} Q(A,\, \mathsf{ALG}(A^*))\,. \tag{3}$$

Two design choices are essential for correctness and fairness. **(i)**Evaluation on the original graph. The partition $\mathcal{P}^* = \mathsf{ALG}(A^*)$ is always evaluated on A, not on A^*. This ensures that gains are measured on the actual network rather than on the transformed input. **(ii)**Base case included. The point $(\alpha_1, \alpha_2) = (1, 0)$ recovers the standard algorithm on the original graph, so $Q^* \geq Q(A, \mathcal{P})$ is guaranteed: MCG-CDP can never perform worse than the baseline.

The weights are explored over a uniform grid with step $\delta = 0.01$, covering all 5,151 valid combinations of (α_1, α_2). To ensure reproducibility, the random seed is fixed before each algorithm call.

3 Benchmark Evaluation

3.1 Experimental Setup

The benchmark in [3] comprises 150 networks: 10 well-known networks (including Zachary Karate Club, Dolphins, Les Misérables, PolBooks, and Football), 80 networks from the Konect repository, and 60 synthetically generated networks (small-world, hierarchical, and scale-free). Five community detection algorithms are used as base algorithms: Louvain, Leiden, Walktrap, Infomap, and Fast Greedy, yielding 750 network–algorithm experiments. For each pair, an experiment is counted as an improvement whenever $Q^* > Q(A, \mathcal{P})$.

3.2 Main Results

MCG-CDP improves modularity in over 76% of the 750 experiments, consistently across all five base algorithms. Table 1 reports the improvement rate $\hat{p}$ and exact 95% confidence intervals for each algorithm [3]. Louvain benefits most ($\hat{p} = 0.800$), while Fast Greedy shows the lowest gain ($\hat{p} = 0.607$). The aggregate 95% CI $[0.704, 0.838]$ lies entirely above 0.5, confirming that improvement significantly exceeds chance across all methods. Networks with already high base modularity are not degraded, since the base case $(\alpha_1, \alpha_2) = (1, 0)$ is always included in the search.

Table 1. Per-algorithm improvement rates and exact 95% confidence intervals on the 150-network benchmark [3].

Algorithm	$\hat{p}$	95% CI
Louvain	0.800	[0.736, 0.864]
Leiden	0.773	[0.706, 0.840]
Walktrap	0.773	[0.706, 0.840]
Infomap	0.760	[0.692, 0.828]
Fast Greedy	0.607	[0.529, 0.685]
All methods	**0.767**	[**0.704**, **0.838**]

3.3 Weight Analysis

A supervised machine learning analysis in [3] shows that the weight parameters α_1 and α_2 each account for approximately 33% of the importance in predicting whether a given combination improves modularity, collectively outweighing the choice of base algorithm. Moreover, more than 10% of all tested (α_1, α_2) combinations improve the baseline for any given network, confirming that effective weight configurations are reasonably frequent and the grid search reliably finds them.

4 Conclusions

The MCG-CDP framework provides a principled and broadly applicable enhancement to matrix-based community detection: by aggregating A, A^2, and A^3 with optimised weights, it improves modularity in the majority of real-world networks without modifying the base algorithm or risking regression below its result.

The full methodology, detailed experimental results across 150 networks, dynamic analysis for time-varying networks, and complexity discussion can be found in the published work:

Hernández, R., Gutiérrez, I., Castro, J.: Social Network Analysis: A Novel Paradigm for Improving Community Detection. *Int. J. Comput. Intell. Syst.* **18**, 87 (2025).
`https://doi.org/10.1007/s44196-025-00812-9`

References

1. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment **2008**(10), 10008 (2008)
2. Fortunato, S., Hric, D.: Community detection in networks: A user guide. Physics Reports **659**, 1–44 (2016)
3. Hernández, R., Gutiérrez, I., Castro, J.: Social network analysis: A novel paradigm for improving community detection. International Journal of Computational Intelligence Systems **18**(87) (2025). `https://doi.org/10.1007/s44196-025-00812-9`
4. Lancichinetti, A., Fortunato, S.: Community detection algorithms: A comparative analysis. Physical Review E **80**(5), 056117 (2009)
5. Newman, M.E.J., Girvan, M.: Finding and evaluating community structure in networks. Physical Review E **69**(2), 026113 (2004)
6. Traag, V.A., Waltman, L., van Eck, N.J.: From Louvain to Leiden: Guaranteeing well-connected communities. Scientific Reports **9**(5233) (2019)