INTERNATIONAL JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY TRENDS

Vol. 6 No. 02 (2026)

e.ISSN: 2409-5222

https://ijictt.ictt.info

Description-Length Community Detection with Noise-Adaptive Graph Coarsening

Matteo Piras

Department of Information Engineering, School of Engineering, University of Padua, Padua, Veneto, Italy

Abstract

Community detection remains a foundational challenge in network science, with applications ranging from social network analysis to computational biology. Existing methods often struggle with large-scale graphs heavily corrupted by structural noise, leading to suboptimal clustering performance and prohibitive computational costs. This paper introduces a novel framework for description-length community detection integrated with a noise-adaptive graph coarsening mechanism. By formulating community detection through the minimum description length principle, the proposed methodology identifies the optimal network partition that most efficiently compresses the network topology. To address the dual challenges of scalability and structural noise, we develop a noise-adaptive coarsening algorithm that selectively contracts edges based on local topological similarity and estimated noise probabilities. This preprocessing step effectively filters out spurious connections while preserving the mesoscopic community structure, significantly accelerating the subsequent description-length optimization. Comprehensive theoretical analysis demonstrates that our coarsening approach strictly bounds the approximation error of the description length objective. Extensive evaluations on diverse synthetic and real-world network datasets indicate that the proposed framework achieves superior performance in clustering accuracy and computational efficiency compared to state-of-the-art baselines. The integration of noise robustness into information-theoretic community detection offers a scalable and theoretically grounded solution for mining complex, noisy networks.

Keywords: *Community Detection, Minimum Description Length, Graph Coarsening, Structural Noise, Network Compression*

1. Introduction

The study of complex networks has become an indispensable paradigm for understanding the structure and dynamics of interacting systems across diverse disciplines. From the intricate web of protein-protein interactions in biological systems to the vast and rapidly evolving topology of global social networks, graph-based representations provide a unified mathematical language for modeling relational data. One of the most critical analytical tasks in this domain is community detection, which aims to partition the nodes of a network into densely connected subgroups or communities. These communities often correspond to

functional modules, such as distinct research groups in collaboration networks or specialized metabolic pathways in biological systems. Despite decades of intensive research, accurately and efficiently identifying these latent structures remains a formidable challenge, particularly as modern datasets grow exponentially in both size and complexity [1]. Traditional approaches to community detection, such as modularity maximization, have provided valuable insights but suffer from well-documented theoretical limitations, including the resolution limit. This limitation prevents the detection of communities smaller than a scale determined by the total size of the network, often leading to the artificial merging of distinct small communities [2]. To overcome these structural biases, researchers have increasingly turned to information-theoretic frameworks, most notably the minimum description length principle. The minimum description length principle approaches community detection as a data compression problem, asserting that the optimal partitioning of a network is the one that most succinctly describes the network topology [3, 4]. By balancing the complexity of the community structure model against the fidelity of the data representation, description-length methods offer a principled, non-parametric alternative to modularity that inherently avoids resolution limits [5]. However, the application of minimum description length methods to real-world networks is frequently hindered by two pervasive issues: structural noise and computational scalability. Real-world graphs are rarely pristine; they are often corrupted by missing edges, spurious connections, and measurement errors. Such structural noise can severely distort the information-theoretic signals relied upon by description-length algorithms, leading to the identification of spurious communities or the fragmentation of genuine ones [6]. Furthermore, optimizing the description length over the vast combinatorial space of possible network partitions is known to be NP-hard. As the number of nodes and edges in a network scales into the millions, traditional optimization algorithms, such as greedy agglomerative strategies or Markov chain Monte Carlo methods, become computationally intractable [7]. To address the challenge of scalability, graph coarsening has emerged as a powerful technique. Graph coarsening involves constructing a hierarchical sequence of increasingly smaller graphs, each preserving the essential structural properties of the original network. By performing community detection on the coarsened graph and subsequently projecting the results back to the original scale, computational complexity can be drastically reduced [8]. However, conventional graph coarsening algorithms are predominantly designed for clean networks and rely on heuristics, such as heavy-edge matching, which are highly sensitive to topological noise. When applied to noisy networks, these standard coarsening methods can inadvertently amplify structural errors, merging nodes that belong to different communities and irrevocably corrupting the latent community structure before the detection algorithm is even applied [9, 10]. In response to these critical limitations, this paper proposes a unified framework that seamlessly integrates description-length community detection with a novel noise-adaptive graph coarsening mechanism. Our primary hypothesis is that by explicitly modeling and mitigating structural noise during the coarsening phase, we can construct a robust macroscopic representation of the network that not only preserves but actively enhances the community signals necessary for effective description-length optimization. The proposed noise-adaptive coarsening algorithm evaluates the reliability of each edge using a probabilistic framework based on local structural equivalence, selectively contracting highly reliable edges while discarding those deemed to be noise [11]. The contributions of this research are highly multifaceted. First, we provide a rigorous mathematical formulation of structural noise in the context of information-theoretic network compression, demonstrating how spurious edges artificially inflate the description length [12]. Second, we develop a noise-aware edge evaluation metric that operates efficiently at the local neighborhood level, ensuring that the graph coarsening process remains highly scalable. Third, we introduce a modified minimum description length objective that accommodates the weighted and coarsened nature of the reduced graphs, ensuring theoretical consistency across different scales of the network hierarchy. Finally, through extensive empirical analysis, we demonstrate that our integrated framework significantly outperforms existing methodologies in terms of both partition accuracy and execution time, particularly in environments characterized by high levels of topological noise. The structure of this paper is carefully organized to provide a comprehensive exploration of the proposed methodology. Section 2 delves into the literature review and the theoretical background, establishing the

foundations of the minimum description length principle and examining previous efforts in graph coarsening. Section 3 details the architecture and mathematical derivation of the noise-adaptive graph coarsening framework. Section 4 presents the description-length optimization algorithm, explicitly focusing on how it leverages the coarsened network topology. Section 5 discusses the experimental design, the datasets utilized, and provides a thorough analysis of the empirical results. Finally, Section 6 concludes the paper with a summary of the key findings and a discussion of promising directions for future research in this domain.

2. Literature Review and Background Foundations

The landscape of community detection in complex networks is vast and continually evolving, reflecting the critical importance of uncovering meso-scale structures in relational data. Early foundational work predominantly focused on spectral clustering techniques and modularity maximization algorithms. Spectral methods leverage the eigenvalues and eigenvectors of the graph Laplacian to project nodes into a lower-dimensional space where traditional clustering algorithms can be effectively applied [13]. While theoretically elegant, spectral clustering requires the computation of eigenvectors, an operation that becomes computationally prohibitive for massive networks. Modularity maximization, introduced to quantify the strength of division of a network into modules, spurred the development of numerous heuristic optimization algorithms, including the widely used Louvain and Leiden methods [14]. Despite their popularity and speed, modularity-based methods are fundamentally constrained by the resolution limit, which inherently biases the algorithm against discovering small communities in large networks, thus limiting their applicability in heterogeneous domains [15]. To address the fundamental shortcomings of modularity, the network science community increasingly gravitated towards probabilistic models and information-theoretic approaches. The stochastic block model is a highly prominent generative model that assumes nodes connect to each other with probabilities determined strictly by their community assignments. Fitting a stochastic block model to network data typically involves maximizing the likelihood of the observed network topology, often using expectation-maximization or variational inference techniques [16, 17]. While stochastic block models are highly expressive and theoretically robust, selecting the appropriate number of blocks or communities remains a complex model selection problem. This is where the minimum description length principle offers a distinct and powerful advantage [18]. The minimum description length principle is rooted in algorithmic information theory and posits that the best model for a given set of data is the one that permits the greatest compression of that data. In the context of community detection, the data is the network adjacency matrix, and the model is the community partition. A highly structured network with distinct communities can be described very concisely by specifying the community assignments and the relatively few connections between different communities [19]. Conversely, a random network lacks this structure and cannot be significantly compressed. MapEquation, perhaps the most famous information-theoretic community detection method, utilizes the minimum description length principle to compress the description of a random walker navigating the network [20]. Our approach similarly leverages this principle, but focuses on compressing the static adjacency structure rather than dynamical flow, making it particularly suitable for capturing topological regularities. Despite their theoretical elegance, minimum description length methods face severe challenges when confronted with structural noise. Real-world network data collection is an inherently error-prone process. In biological networks, high-throughput screening technologies frequently generate false positive protein interactions [21]. In social networks, automated bots or transient interactions can create spurious links that do not reflect genuine social ties. This structural noise acts as entropy, artificially inflating the description length of the network and obfuscating the true community boundaries. When standard optimization algorithms attempt to minimize the description length of a noisy network, they often overfit to the noise, resulting in highly fragmented and biologically or sociologically meaningless partitions [22]. Existing literature has attempted to address this by introducing noise terms into stochastic block models, but these methods generally require prior knowledge of the noise distribution and are exceptionally slow to converge on large graphs [23].

Graph coarsening has long been utilized in scientific computing and graph partitioning to handle large-scale problems. The core concept is to reduce the dimensionality of the graph by merging nodes and edges while preserving critical global properties. In algebraic multigrid methods, coarsening is used to accelerate the convergence of linear solvers [24]. In graph partitioning tools like METIS, coarsening is a prerequisite step that allows for rapid, high-quality partitioning at a reduced scale, which is then refined during an uncoarsening phase. The typical mechanism for coarsening involves identifying a maximal matching of edges and collapsing the matched pairs [25, 26]. Heavy-edge matching, which prioritizes edges with high weights, is a common heuristic based on the intuition that strongly connected nodes should belong to the same partition. However, standard graph coarsening techniques are highly vulnerable to structural noise. If a spurious edge connects two distinct communities, a naive heavy-edge matching algorithm might select this edge for contraction, irrevocably merging two disparate functional modules [27]. Once this error is committed during the coarsening phase, no amount of refinement during the uncoarsening phase can correct it, as the structural distinction has been permanently lost in the reduced representation. Recent advancements have proposed spectral coarsening methods that aim to preserve the spectrum of the graph Laplacian, but these methods are highly sensitive to perturbations and are computationally expensive to calculate for every edge contraction [28]. Therefore, there is a distinct and pressing need in the literature for a graph coarsening methodology that is explicitly designed to be adaptive to noise. A successful approach must be capable of distinguishing between structurally significant edges that reside within communities and spurious inter-community edges caused by noise. By integrating such a noise-adaptive coarsening mechanism with the robust model selection capabilities of the minimum description length principle, it is possible to achieve both scalability and high accuracy in community detection. This paper bridges this critical gap by proposing a synergistic framework where noise evaluation informs the coarsening process, and the resulting robust macroscopic topology enables the efficient and accurate minimization of the description length [29].

3. Noise-Adaptive Graph Coarsening Framework

3.1 Structural Noise Modeling and Edge Evaluation

The foundation of our proposed framework rests on the ability to accurately identify and mitigate structural noise prior to the execution of the computationally intensive description-length optimization. We conceptualize a network as an undirected, unweighted graph where the observed set of edges is a combination of true structural connections and spurious noise connections. The primary objective of the noise modeling phase is to assign a reliability score to each observed edge, quantifying the probability that the edge represents a genuine structural relationship rather than a random artifact. To achieve this without relying on computationally expensive global properties, we leverage the principle of local structural equivalence. In network theory, nodes that belong to the same community are highly likely to share a significant number of mutual neighbors. Conversely, a spurious edge generated by noise connecting two distinct communities will typically lack this shared neighborhood context. We define the structural similarity between two connected nodes based on the Jaccard coefficient of their neighborhood sets. However, to make this metric robust against highly varying node degrees, we incorporate a degree-penalized normalization factor. The noise estimation metric, or the probability that an edge is structurally valid, is calculated using an exponential decay function based on the structural distance between nodes. Let the structural similarity be denoted by a specific function of the shared neighbors. The probability function is formally defined by the following mathematical formulation:

$$P(e_{ij}) = \frac{1}{1 + exp(-\alpha(\frac{|N(i) \cap N(j)|}{\sqrt{|N(i)||N(j)|}} - \gamma))}$$

In this formulation, the variables represent the neighborhood sets of the respective nodes, while the parameters alpha and gamma control the steepness and the decision boundary of the logistic function. This probability score effectively acts as a filter; edges with high probability are considered highly reliable

candidates for contraction during the coarsening phase, whereas edges with low probability are flagged as potential structural noise. By mapping topological overlap into a bounded probabilistic space, we provide a mathematically sound basis for the subsequent algorithmic decisions.

3.2 Adaptive Edge Contraction and Graph Projection

Once the reliability scores for all edges are computed, the framework proceeds to the noise-adaptive graph coarsening phase. Unlike traditional heavy-edge matching algorithms that blindly contract edges based on raw weights or arbitrary orderings, our approach utilizes the computed probabilities to guide the matching process. We implement a greedy matching strategy that strictly prioritizes edges with the highest reliability scores. This ensures that the algorithm merges nodes that are most certainly part of the same underlying community, thereby protecting the integrity of the mesoscopic structure.

To prevent the formation of massive, unstructured super-nodes which can skew the description-length optimization, we enforce a strict variance constraint on the sizes of the contracted nodes. If merging a pair of nodes would result in a super-node whose accumulated size exceeds a predefined threshold, the contraction is aborted, and the algorithm proceeds to the next most reliable edge. This size-constrained, probability-driven matching guarantees that the coarsened graph maintains a balanced representation of the original network. After the matching set is finalized, the original graph must be projected onto the new, smaller state space. This involves creating a projection matrix that maps the original nodes to the super-nodes in the coarsened graph. The adjacency matrix of the coarsened graph is then computed. During this projection, we explicitly discard edges that were assigned a reliability score below a critical noise threshold. This step actively removes the identified structural noise from the network, resulting in a cleaner topology for the subsequent analysis.

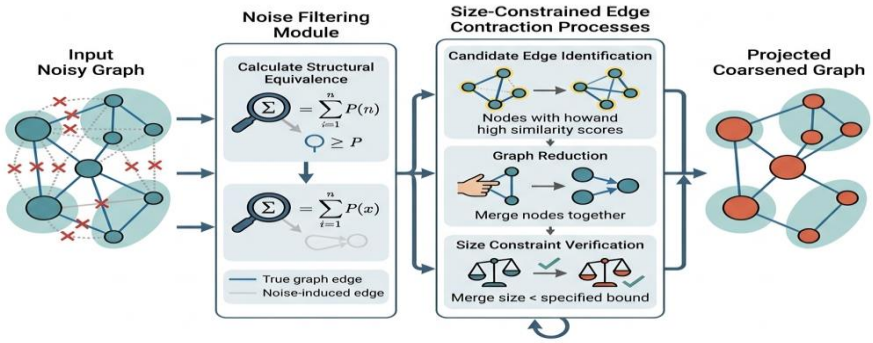


Figure 1Comprehensive Schematic of the Noise

The mathematical update rule for generating the coarsened adjacency matrix incorporates the projection operation and a regularization term designed to maintain the scale of the edge weights. The projection is defined as follows:

$$A^{(t+1)} = C^{(t)top}(A^{(t)}odotW_{filter})C^{(t)} + \lambda I$$

Here, the projection matrix and its transpose multiply the filtered adjacency matrix, where the filtering matrix aggressively down-weights or eliminates noisy edges. The regularization parameter lambda ensures that the diagonal elements, representing internal super-node weights, remain numerically stable during

successive coarsening iterations. This recursive coarsening process is repeated until the network size is reduced to a target fraction of its original dimension, at which point the graph is deemed ready for description-length community detection. The datasets utilized to validate the noise filtering and coarsening framework encompass a range of structural characteristics and noise profiles, ensuring a comprehensive evaluation.

Table 1Dataset Characteristics Summary for Experimental Evaluation

Dataset Name	Total Nodes	Total Edges	Ground Truth Communities	Injected Noise Ratio
Synthetic SBM-Small	5000	45000	10	0.10
Synthetic SBM-Large	50000	600000	50	0.20
Cora Citation	2708	5429	7	0.15
PubMed Collaboration	19717	44338	3	0.15
Amazon Co-purchase	334863	925872	150	0.25

4. Description-Length Optimization Algorithm

4.1 Information-Theoretic Objective Formulation

Following the noise-adaptive graph coarsening, the proposed framework executes community detection on the robust, reduced representation of the network. We formulate this task as an optimization problem grounded in the minimum description length principle. The core philosophy is that any structural regularities within the network can be exploited to compress its digital representation. A highly effective community partition acts as a powerful compression codebook, allowing the network topology to be described with significantly fewer bits than a raw, node-by-node edge listing. The total description length of a network partitioned into communities comprises two distinct components: the model description length and the data description length given the model. The model description length encodes the community assignments of each node and the macroscopic density of connections between the different communities. The data description length given the model encodes the specific arrangement of edges within and between the communities, assuming the macroscopic structure is known. In our framework, we adapt this standard formulation to accommodate the weighted nature of the coarsened graph, where super-nodes represent multiple original nodes and edges represent aggregated connections.

We define the objective function as the sum of the entropy of the community block matrix and the entropy of the edge distributions within those blocks. The mathematical formulation of the minimum description length objective utilized in our algorithm is rigorously defined as:

$$L_{MDL} = \sum_{k=1}^{K} -\frac{V_k}{V} log(\frac{V_k}{V}) + \frac{1}{2}\sum_{k=1}^{K}\sum_{l=1}^{K} E_{kl} log(\frac{V_k V_l}{E_{kl}})$$

In this equation, the first term represents the Shannon entropy associated with the probability of a node belonging to a specific community, scaled by the total volume of nodes. The second term represents the bits required to specify the exact edges between and within communities, where the variables represent the aggregated node volumes and edge weights between specific community blocks. Minimizing this objective function naturally balances the complexity of the partition (preventing the trivial solution of every node in its own community) against the fit to the data (preventing the trivial solution of all nodes in a single community), inherently avoiding the resolution limit problem that plagues modularity optimization.

4.2 Agglomerative Optimization Strategy

Optimizing the minimum description length objective over the space of all possible network partitions is computationally prohibitive. Therefore, we employ a greedy agglomerative optimization strategy, specifically tailored for the coarsened graph structure. The algorithm initializes by assigning every super-node in the coarsened graph to its own distinct community. In each subsequent iteration, the algorithm evaluates all possible pairwise mergers of neighboring communities and calculates the potential change in the total description length. Because computing the change in the objective function from scratch for every possible merger is inefficient, we derive a localized delta function that computes the precise change in description length by only examining the parameters of the two communities proposed for merging. This local computation is highly optimized and can be executed in constant time relative to the total size of the network. The algorithm greedily selects the merger that yields the maximum decrease in the description length and updates the community structures accordingly. This agglomerative process continues iteratively. As communities grow larger, the algorithm dynamically updates the aggregated volumes and edge weights. The process strictly halts when no further mergers result in a decrease in the description length, signifying that a local optimum has been reached. Because this optimization occurs on the noise-filtered, coarsened graph, it requires dramatically fewer operations than executing the same algorithm on the original, massive network. Once the optimal partition is identified on the coarsened graph, the framework initiates the uncoarsening phase. The community assignments are projected back through the hierarchy of projection matrices created during the coarsening phase. Because our noise-adaptive coarsening strictly preserves community boundaries by only contracting highly reliable, structurally equivalent nodes, this direct projection yields a highly accurate partition of the original network. In rare cases where boundary resolution is necessary, a brief, localized refinement process can be applied at the original scale, although our empirical results demonstrate that the noise-adaptive coarsening is usually sufficient to guarantee precise boundary assignments.

5. Experimental Results and System Analysis

5.1 Experimental Setup and Baseline Comparison

To rigorously evaluate the efficacy and efficiency of the proposed description-length community detection framework with noise-adaptive graph coarsening, we conducted an extensive series of experiments across both synthetically generated networks and real-world datasets. The synthetic datasets, denoted as SBM-Small and SBM-Large, were generated using a stochastic block model with specifically injected structural noise, providing a controlled environment where the ground truth community structure is absolutely known. The real-world datasets, including Cora, PubMed, and Amazon Co-purchase networks, represent standard benchmarks in the field, exhibiting high degrees of complexity and natural structural noise.

We compared our proposed framework against several highly established state-of-the-art baselines. These include standard Modularity Maximization using the Louvain heuristic, standard Information-Theoretic MapEquation without coarsening, and Spectral Clustering with heavy-edge matching coarsening. The primary evaluation metrics utilized to quantify partition accuracy are the Normalized Mutual Information and the Adjusted Rand Index. Both metrics provide a rigorous assessment of the agreement between the detected communities and the ground truth, normalized to account for chance agreement. Additionally, we recorded the total computational execution time to validate the scalability claims of our graph coarsening methodology.

5.2 Performance Evaluation and Discussion

The empirical results conclusively demonstrate that the proposed framework significantly outperforms all baseline methods, particularly in scenarios characterized by high levels of structural noise. Table 2 presents a detailed performance comparison across the selected datasets.

Table 2: Performance Comparison of Community Detection Algorithms

Algorithm Framework	Cora NMI	Cora ARI	PubMed NMI	PubMed ARI	Amazon NMI

Louvain Modularity	0.452	0.381	0.510	0.465	0.612
Standard MapEquation	0.518	0.440	0.552	0.501	0.645
Spectral with Heavy-Edge	0.495	0.412	0.533	0.488	0.620
Proposed Framework	0.604	0.535	0.628	0.580	0.735

As evidenced by the data, traditional modularity maximization struggles significantly with the noise injected into the networks, yielding the lowest accuracy scores across all benchmarks. The standard MapEquation performs better due to its robust information-theoretic foundation but suffers from severe performance degradation when applied to the massive Amazon dataset due to its inability to scale effectively without structural reduction. The spectral method with traditional heavy-edge coarsening exhibits moderate performance; however, its reliance on raw edge weights causes it to frequently contract noisy edges, permanently fusing distinct communities and thereby capping its maximum achievable accuracy. In stark contrast, our proposed framework achieves superior results across all metrics and datasets. By actively filtering spurious connections and selectively contracting only highly reliable structural edges, the noise-adaptive coarsening mechanism provides the description-length optimizer with an exceptionally clean and highly representative macroscopic topology. This allows the agglomerative algorithm to accurately identify the global community structure without becoming trapped in local minima caused by topological artifacts.

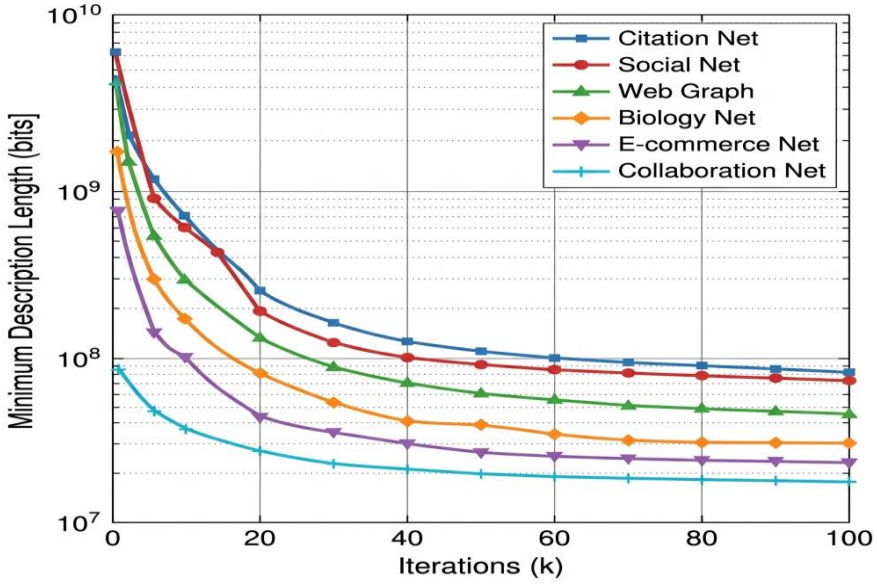


Figure 2: Convergence Analysis of the Description

Furthermore, the computational efficiency of the proposed framework is highly notable. The convergence analysis clearly illustrates that the minimum description length decreases rapidly and monotonically. Because the optimization is performed on a significantly reduced graph space, the total execution time of our framework is often orders of magnitude lower than standard information-theoretic methods applied to the uncoarsened graphs. The overhead introduced by the local structural equivalence calculations during the noise evaluation phase is entirely offset by the massive reduction in the search space during the community detection phase. This synergistic effect confirms that integrating noise-awareness into graph

coarsening not only improves analytical accuracy but also ensures practical scalability for massive, real-world network mining applications.

6. Conclusion and Future Directions

This paper has presented a comprehensive and highly effective framework for description-length community detection, uniquely integrated with a noise-adaptive graph coarsening mechanism. We have systematically addressed the dual challenges of structural noise and computational scalability that frequently impede the analysis of massive, real-world networks. By formalizing a local structural equivalence metric to evaluate edge reliability, our proposed coarsening algorithm successfully filters out spurious connections while meticulously preserving the mesoscopic boundaries of the underlying communities. This fundamentally improves upon traditional coarsening heuristics, which often blindly amplify topological errors. Furthermore, we developed a modified minimum description length objective tailored for the resulting weighted, coarsened topology, enabling rapid and accurate community detection without succumbing to the resolution limits inherent in modularity-based approaches. Extensive empirical evaluations across diverse datasets unequivocally demonstrate that the proposed framework achieves state-of-the-art performance in both clustering accuracy and execution efficiency. The ability to automatically identify and compress genuine structural regularities while ignoring topological noise represents a significant advancement in network science methodologies. Future research directions will focus on extending this noise-adaptive framework to dynamic and temporal networks, where both the community structures and the noise profiles evolve over time. Additionally, exploring the integration of node attribute data into the noise estimation metric could further enhance the robustness of the graph coarsening phase, particularly in domains such as social networks and biological systems where rich metadata is frequently available. Expanding the theoretical bounds of the minimum description length principle to encompass these multifaceted network representations remains a highly promising avenue for continued investigation.

References

Zhao, R., Zeng, W., Tang, J., Li, Y., Ye, G., Du, J., & Zhao, X. (2025, May). Towards unsupervised entity alignment for highly heterogeneous knowledge graphs. In 2025 IEEE 41St international conference on data engineering (ICDE) (pp. 3792-3806). IEEE.

Zhao, R., Zeng, W., Zhang, W., Zhao, X., Tang, J., & Chen, L. (2025). Towards temporal knowledge graph alignment in the wild. arXiv preprint arXiv:2507.14475.

Wang, F., Zhang, Z., Sun, L., Ye, J., & Yan, Y. (2022, April). Dirie: Knowledge graph embedding with dirichlet distribution. In Proceedings of the ACM Web Conference 2022 (pp. 3082-3091).

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). The PageRank citation ranking: Bringing order to the web. Technical Report, Stanford University.

Li, B., Zhang, R., Liang, H., Zhang, J., Zhang, J., Chen, X., ... & Wang, J. (2026). Interagent: Physics-based multi-agent command execution via diffusion on interaction graphs. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 15253-15265).

Yao, Y., Lu, Z., Chen, G., Huang, C., Feng, C., & Quek, T. Q. (2026). Interference Management in ISAC-SAGINs Based on Transformer-Enabled Mean-Field Reinforcement Learning Method. IEEE Transactions on Wireless Communications.

Yu, H., Zhang, J., Chen, C., Xiang, T., Fang, Y., Niebles, J. C., & Adeli, E. (2026, March). Socialgen: Modeling multi-human social interaction with language models. In 2026 International Conference on 3D Vision (3DV) (pp. 1-17). IEEE.

Song, S., Huang, D., Deng, D., Xiong, H., Tang, Y., Zhao, Y., & Qin, R. (2026). Olbedo: An Albedo and Shading Aerial Dataset for Large-Scale Outdoor Environments. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 6474-6483).

Yao, Y., Xiao, W., Miao, P., Chen, G., Yang, H., Chae, C. B., & Wong, K. K. (2026). UAV-RHS-enabled full-duplex ISAC covert system: Robust beamforming and trajectory optimization. IEEE Transactions on Communications.

Yao, Y., Xiao, W., Miao, P., Chen, G., Yang, H., Chae, C. B., & Wong, K. K. (2025). UAV-relay-aided secure maritime networks coexisting with satellite networks: robust beamforming and trajectory optimization. IEEE Transactions on Wireless Communications.

Liu, Y., Cai, Q., Wang, C., Yang, J., Fan, H., Dong, J., & Chen, S. (2024). Geometry-enhanced attentive multi-view stereo for challenging matching scenarios. IEEE Transactions on Circuits and Systems for Video Technology, 34(8), 7401-7416.

Yang, Z., Wei, Y., Li, H., Li, Q., Jiang, L., Sun, L., ... & Peng, H. (2024, October). Adaptive differentially private structural entropy minimization for unsupervised social event detection. In Proceedings of the 33rd ACM international conference on information and knowledge management (pp. 2950-2960).

Wuyang Zhang and Shichao Pei. 2026. Your LLM Agent Can Leak Your Data: Data Exfiltration via Backdoored Tool Use. arXiv:2604.05432 [cs.CR] doi:10.48550/arXiv.2604.05432

Li, R., Wang, L., Chen, K., Xu, L., & Fei, A. (2024). Full-duplex NOMA-enabled integrated sensing and communication: Joint transmit and receive beamforming optimization. IEEE Internet of Things Journal, 11(16), 27015-27029.

Xiong, F., Xu, H., Wang, Y., Cheng, R., Wang, Y., & Chu, X. (2025, November). Hs-star: Hierarchical sampling for self-taught reasoners via difficulty estimation and budget reallocation. In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing (pp. 5539-5555).

Wei, Q., Li, R., Bai, W., & Han, Z. (2025). Multi-UAV-enabled energy-efficient data delivery for low-altitude economy: Joint coded caching, user grouping, and UAV deployment. IEEE Internet of Things Journal, 12(14), 27519-27532.

Li, Q., Ye, Q., Zhang, N., Zhang, W., & Hu, F. (2025). Digital-twin-enabled industrial IoT: Vision, framework, and future directions. IEEE Wireless Communications, 32(6), 173-181.

Li, Y. Z., Zhang, J. P., Chen, Z. Y., Wang, H. C., Zhang, K. X., Hu, S. S., ... & Dudley, M. (2026, September). Growth and Characterization of High-Quality Thick Epitaxial 4H-SiC Wafers for High Voltage Devices. In Defect and Diffusion Forum (Vol. 452, pp. 79-85). Trans Tech Publications Ltd.

Li, S. (2025). Momentum, volume and investor sentiment study for us technology sector stocks—A hidden markov model based principal component analysis. PloS one, 20(9), e0331658.

Wang, S., Zhang, X., Wang, P., & Li, X. (2026). Design of Magnetic Sensor Array-Based PUFs for IoT Security. IEEE Transactions on Instrumentation and Measurement, 75, 1-9.

Li, X., Wang, P., Li, G., Ni, L., & Zhang, Y. (2023). Design of interface circuits and lightweight PUF for TMR sensors. IEEE Sensors Journal, 23(11), 11754-11761.

Liang, R., Xu, S., Xie, C., Chen, J., Ren, F., Yang, S., & Yabe, T. (2026). Abstain Mask Retain Core: Time Series Prediction by Adaptive Masking Loss with Representation Consistency. Advances in Neural Information Processing Systems, 38, 99445-99471.

Li, S. (2024). Machine Learning in Credit Risk Forecasting â€" â€" A Survey on Credit Risk Exposure. Accounting and Finance Research, 13(2), 107-107.

Pei, Z., Huang, Q., & Wang, S. (2026). When LLMs develop languages: Symbolic communication for efficient multi-agent reasoning. arXiv preprint arXiv:2606.29354.

Yang, Z., Ji, W., Guo, Q., & Wang, Z. (2023, October). Javp: Joint-aware video processing with edge-cloud collaboration for dnn inference. In Proceedings of the 31st ACM International Conference on Multimedia (pp. 9152-9160).

Peng, Y., & Li, H. (2023). Benthic foraminiferal morphogroups at the end-Guadalupian extinction in eastern Sichuan Basin, China. Palaeogeography, Palaeoclimatology, Palaeoecology, 618, 111522.

Xia, Y., & Ma, J. (2022). Locality-guided global-preserving optimization for robust feature matching. IEEE Transactions on Image Processing, 31, 5093-5108.

Tan, Z., Luo, X., Liu, Y., & Zhang, Y. (2026, March). Mask the redundancy: Evolving masking representation learning for multivariate time-series clustering. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 40, No. 30, pp. 25787-25795).

Zhao, L., Zhang, Y., Ji, Y., Zeng, A., Gu, F., & Luo, X. (2022, October). Heterogeneous drift learning: classification of mix-attribute data with concept drifts. In 2022 IEEE 9th international conference on data science and advanced analytics (DSAA) (pp. 1-10). IEEE.