Algorithms and Intelligence

Vol 2 Issue 1 (2026)

www.algointel.org

Stability-Regularized Temporal Graph Learning for Dynamic Community Detection in Evolving Networks

Xiao Yu

School of Data Science, Lingnan University, Hong Kong 999077, China

Abstract

The study of evolving networks has become increasingly crucial for understanding complex systems ranging from social media interactions to biological neural pathways. While standard community detection algorithms excel in static environments, they frequently struggle with dynamic graphs, often producing highly unstable community partitions that fluctuate erratically due to minor structural perturbations. This paper introduces a comprehensive framework for stability-regularized temporal graph learning, specifically designed to mitigate the volatility inherent in dynamic community detection. By integrating a temporal regularization term directly into the optimization landscape of a graph neural network, the proposed architecture enforces temporal smoothness while retaining sufficient plasticity to adapt to genuine structural evolutions. The methodology involves processing continuous network snapshots through a specialized temporal attention mechanism, balanced by a novel stability loss function that penalizes unwarranted deviations in cluster assignments across consecutive time steps. Extensive empirical evaluations are conducted across multiple distinct datasets, representing communication networks, citation graphs, and transactional environments. The results demonstrate that the stability-regularized approach significantly outperforms existing state-of-the-art dynamic graph learning models in both clustering accuracy and temporal consistency metrics. The framework offers a robust theoretical and practical foundation for analyzing highly volatile network topologies, providing deeper insights into the underlying mechanisms of structural evolution in complex domains.

Keywords

Temporal Graph Learning, Community Detection, Evolving Networks, Stability Regularization, Graph Neural Networks

1. INTRODUCTION

Background on Dynamic Networks

Complex systems in the real world are rarely static. They are characterized by continuous evolution, where entities and their relationships undergo constant modifications over time. These systems are most effectively modeled as dynamic or temporal networks. In disciplines ranging from computational sociology to bioinformatics, understanding the meso-scale structures of these networks is fundamentally important [1]. Community detection, the process of grouping network nodes into clusters such that internal connections are dense while external connections are sparse, serves as a primary tool for uncovering these hidden functional modules. In static graphs, this problem has been extensively explored, yielding robust solutions based on modularity maximization, spectral clustering, and random walk dynamics [2]. However, when network topologies evolve, applying static community detection algorithms to individual temporal snapshots independently often leads to a severe degradation in longitudinal coherence. Nodes may artificially oscillate between communities across consecutive timestamps due to algorithmic artifacts rather than genuine systemic shifts, a phenomenon that significantly undermines the interpretability of the extracted network dynamics [3]. Consequently, researchers have shifted their focus toward dynamic community detection paradigms that treat the temporal dimension not merely as a sequence of independent states, but as an integral component of the network topology itself [4]. Addressing this challenge requires sophisticated methodologies capable of capturing both the spatial topology of the graph and the temporal dependencies governing its evolution [5].

1.2 Challenges in Temporal Graph Learning

The integration of deep learning with graph-structured data has led to the proliferation of Graph Neural Networks. These architectures have revolutionized node representation learning by aggregating feature information from local neighborhoods. Adapting these static models to temporal graphs introduces profound computational and theoretical challenges [6]. The primary obstacle is the stability-plasticity dilemma. A dynamic community detection model must possess the plasticity to recognize and adapt to newly forming structures, such as the emergence of a new social group or the dissolution of a financial coalition [7]. Simultaneously, the model must maintain structural stability, ensuring that random noise or minor edge insertions and deletions do not trigger catastrophic reorganizations of the overall community partition [8]. Standard temporal graph learning models, which often utilize recurrent neural networks or self-attention mechanisms to process sequences of graph snapshots, typically prioritize predictive accuracy for future link occurrences [9]. When these standard models are repurposed for clustering, they frequently exhibit high variance in their community assignments. Furthermore, the discrete nature of clustering objectives complicates gradient-based optimization in neural architectures [10]. Minor perturbations in the continuous node embedding space can cross decision boundaries, causing entire subgraphs to abruptly switch community labels. This instability obscures the true evolutionary trajectory of the network, making it difficult for domain experts to distinguish between genuine systemic transformations and mere algorithmic noise [11]. Overcoming this

barrier requires explicit regularization mechanisms that constrain the evolutionary trajectory of the node embeddings.

1.3 Proposed Stability-Regularized Approach

To resolve the tension between adaptability and consistency, this paper proposes a novel framework centered on stability-regularized temporal graph learning. The core hypothesis is that authentic community evolution occurs gradually, meaning that a node's structural role and group affiliation at a given time step should be strongly correlated with its state in the immediate past, unless there is overwhelming topological evidence to the contrary [12]. The proposed methodology enforces this principle through a dual-objective optimization strategy. The primary objective is a self-supervised graph modularity loss that drives the formation of dense, cohesive clusters within the current temporal snapshot [13]. The secondary objective is a stability regularization term that explicitly penalizes divergent embedding trajectories and drastic shifts in community assignment probability distributions between consecutive timestamps. By jointly optimizing these objectives, the model learns temporally smoothed node representations that are robust to transient structural noise [14]. The architecture leverages a temporal graph attention mechanism to weigh historical neighborhood features dynamically, allowing the network to retain a memory of past topologies while prioritizing recent structural changes. Through rigorous mathematical formulation and extensive empirical validation, this research demonstrates that explicit stability regularization not only improves the longitudinal coherence of the detected communities but also enhances the overall accuracy of the structural partitioning across highly volatile network environments.

2. Literature Review

2.1 Traditional Community Detection

The problem of community detection has roots in graph theory and statistical mechanics. Early traditional approaches primarily focused on static networks, attempting to optimize global topological metrics. The introduction of modularity provided a quantitative measure to evaluate the quality of a network partition by comparing the density of edges within communities to a null model of random connections [15]. Numerous heuristic algorithms, such as the Louvain method and the Girvan-Newman algorithm, were developed to maximize this metric efficiently. While successful for static structural analysis, extending these traditional methods to evolving networks initially involved simple two-step procedures. Researchers would either aggregate the dynamic network over a specific time window into a single static graph or apply static detection algorithms to each temporal snapshot independently [16]. Both strategies possess significant flaws. Aggregation destroys temporal resolution, obscuring the precise timing of community formations and splits. Conversely, independent snapshot processing suffers from severe instability, where communities exhibit high volatility, making tracking them over time practically impossible [17]. Evolutionary clustering frameworks were subsequently introduced to address this, incorporating an explicit cost function that penalized deviations from the previous timestamp's partition. These early evolutionary models laid the conceptual groundwork for temporal smoothness but lacked the

capacity to learn complex, non-linear node representations necessary for high-dimensional or attribute-rich networks [18].

2.2 Deep Learning for Temporal Graphs

The advent of representation learning shifted the paradigm from discrete algorithmic optimization to continuous embedding spaces. Graph Neural Networks emerged as powerful tools capable of mapping network nodes into low-dimensional vectors while preserving structural properties. For temporal networks, researchers combined these graph convolutional architectures with sequence models like Long Short-Term Memory networks or Gated Recurrent Units [19]. In these temporal graph network models, a static graph neural network extracts spatial features from a given snapshot, and the sequence model updates a hidden state vector for each node over time. More recent architectures have adopted continuous-time frameworks, treating network evolution as a sequence of asynchronous events rather than discrete snapshots [20]. These event-based models update node embeddings instantaneously upon the arrival of new edges, utilizing temporal attention mechanisms to weigh the relevance of historical interactions. While highly effective for dynamic link prediction and node classification, applying these representation learning frameworks directly to community detection remains problematic [21]. The continuous embedding spaces do not inherently enforce modular structures, and extracting discrete community partitions typically requires a secondary clustering step, such as applying k-means to the learned node vectors. This decoupling of the representation learning phase from the clustering objective often results in sub-optimal community boundaries that fail to capture the true meso-scale dynamics of the network [22].

2.3 Stability and Regularization Techniques

To bridge the gap between deep representation learning and dynamic community detection, recent studies have begun exploring integrated optimization strategies. A critical area of focus is the development of temporal regularization techniques that constrain the learning process. Some approaches involve adding smoothing penalties to the node embeddings, ensuring that the Euclidean distance between a node's vector representation at consecutive timestamps remains small [23]. Other methods incorporate temporal graph regularization directly into the loss function, penalizing structural discrepancies between adjacent snapshots. However, simple embedding smoothing can lead to over-regularization, preventing the model from adapting to genuine sudden shifts in the network topology [24]. To achieve a more nuanced balance, researchers have investigated probabilistic formulations, modeling community assignments as dynamic distributions [25]. Despite these advancements, a persistent challenge remains the precise formulation of the stability penalty. Many existing models apply a uniform penalty across all nodes, disregarding the fact that some nodes inherently exhibit more volatile behavior, such as spanning boundaries between multiple communities, while others represent the stable core of a single module [26]. The methodology presented in this paper addresses these limitations by introducing an adaptive stability regularization term that dynamically scales the temporal penalty based on the structural context and historical volatility of each individual node, thereby preserving the natural heterogeneity of network evolution [27].

3. Methodology and Problem Formulation

3.1 Network Definitions and Temporal Dynamics

We define an evolving network as a sequence of discrete graph snapshots recorded over consecutive time intervals. Let the dynamic network be represented as a temporally ordered set of graphs, where each snapshot captures the structural state of the system at a specific timestamp. Each snapshot consists of a set of nodes and a set of edges connecting them, which may represent interactions, communications, or physical links depending on the specific domain. The adjacency matrix for a given timestamp encodes the presence or absence of edges, and in many real-world scenarios, it can also encompass edge weights denoting the strength of the connections. Additionally, nodes may possess associated feature vectors, such as user profiles in a social network or semantic text in a citation graph. The objective of dynamic community detection is to map the nodes at each timestamp into a set of discrete, non-overlapping or overlapping clusters. This mapping must accurately reflect the underlying density of connections at the current time while maintaining logical consistency with the historical trajectory of the community structures. The fundamental challenge involves tracking the evolution of these communities, including their formation, expansion, contraction, merging, splitting, and eventual dissolution, without allowing algorithmic artifacts to distort the sequence of events.

3.2 Stability-Regularized Objective Function

To achieve robust dynamic community detection, we formulate a joint optimization problem that unifies structural clustering with temporal stability. We project the nodes into a continuous latent space and apply a soft cluster assignment mechanism. The total loss function consists of two primary components. The first component is the structural clustering loss, which encourages the formation of cohesive groups by maximizing the likelihood of observing the current network topology given the cluster assignments. The second component is the stability regularization term, which penalizes abrupt changes in the cluster assignment distributions across successive timestamps. The formulation of this combined objective is mathematically defined as follows:

$$L_{total} = \sum_{t=1}^{T}(L_{struct}^{(t)} + \lambda \sum_{i=1}^{N} D_{KL}(P_i^{(t)} | P_i^{(t-1)}))$$

In this equation, the total loss over all timestamps integrates the structural loss for the specific snapshot and a temporal penalty. The stability term utilizes the Kullback-Leibler divergence to measure the discrepancy between a node's community assignment probability distribution at the current timestamp and its distribution at the previous timestamp. The hyperparameter lambda controls the strength of this regularization, explicitly dictating the trade-off between allowing structural plasticity and enforcing temporal smoothness. By minimizing this divergence, the model is coerced into seeking community assignments that not only fit the immediate topological data but also align closely with the established historical context of each node.

3.3 Temporal Graph Neural Network Architecture

The mechanism for computing the node representations relies on a specialized Temporal Graph Neural Network architecture. This network applies graph convolutional layers to aggregate localized structural information, combined with a temporal attention module that processes the sequence of a node's past states. At each timestamp, the graph convolution computes an intermediate spatial representation by passing messages along the edges present in the current adjacency matrix. Subsequently, the temporal module updates the final node embedding by computing a weighted combination of this spatial representation and the node's historical embeddings. This dynamic aggregation process is formalized by the following layer propagation rule:

$$H^{(t)}_{(l+1)} = \sigma(\alpha^{(t)}(hatA^{(t)}H^{(t)}_{(l)}W_{spatial}) + (1 - \alpha^{(t)})(H^{(t-1)}_{(l+1)}W_{temporal}))$$

Here, the hidden representation for the next layer relies on the normalized adjacency matrix multiplying the current layer's features, transformed by a spatial weight matrix. This is combined with the historical hidden state transformed by a temporal weight matrix. The balancing factor alpha acts as an attention mechanism, determining how much the model should rely on the fresh spatial topology versus the established temporal memory. The non-linear activation function allows the network to learn complex, non-linear structural patterns. This architecture ensures that the node embeddings inherently possess temporal context before the cluster assignments are even calculated, providing a highly stable foundation for the subsequent stability-regularized optimization phase.

4. Algorithmic Design and Implementation

4.1 Optimization Strategy

Optimizing the proposed neural architecture requires a careful approach to handle the dependencies across time. Because the loss function at a given timestamp depends on the model's outputs from previous timestamps, standard independent training procedures are insufficient. We employ a variant of backpropagation through time to optimize the model parameters over sequential windows of network snapshots. To prevent the vanishing gradient problem common in long sequences, we utilize a truncated backpropagation approach, maintaining a fixed-size historical buffer of the previous cluster assignment distributions. The parameter update rule during the optimization process applies gradient descent with respect to the total stability-regularized loss function. The parameter adjustments can be expressed as:

$$\theta_{new} = \theta_{old} - eta\nabla_{\theta}(\frac{1}{B}\sum_{b=1}^{B}L_{total}^{(b)} + \gamma|\theta|_2^2)$$

This formulation demonstrates that the new parameter values are determined by stepping in the opposite direction of the gradient of the batched total loss, scaled by the learning rate eta. An additional L2 regularization term, scaled by gamma, is included to prevent overfitting the model weights to specific topological anomalies present in the training sequence. To handle the discrete

nature of community tracking, the soft cluster assignments are eventually hardened using an argmax operation during the inference phase. We also implement a dynamic learning rate schedule that gradually decays the step size, allowing the model to make significant initial adjustments to the embedding space and subsequently fine-tune the community boundaries for maximal stability.

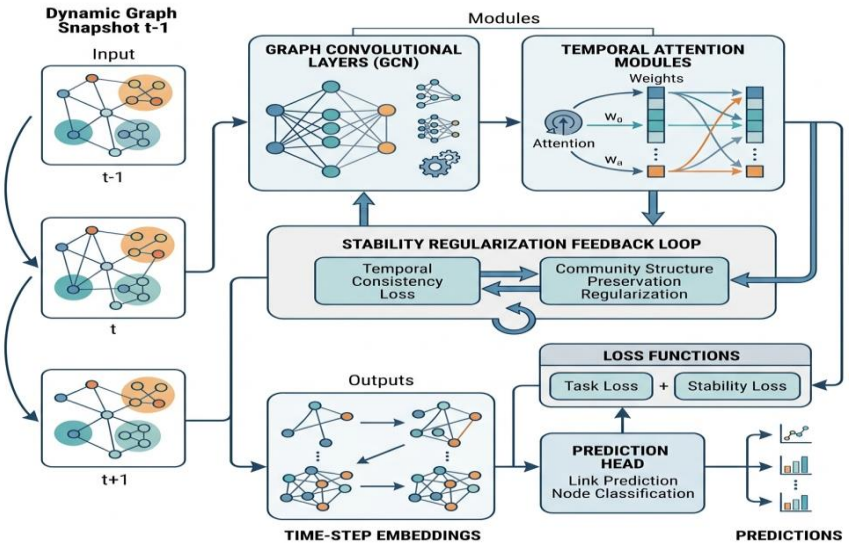


Figure 1 Comprehensive Schematic of the Stability

4.2 Computational Complexity Analysis

The implementation of dynamic network algorithms must account for severe scalability constraints, as real-world graphs frequently comprise millions of nodes and edges. The computational complexity of the proposed framework is primarily governed by the graph convolutional operations and the calculation of the stability regularization term. For a single temporal snapshot containing a specific number of nodes and edges, the sparse matrix multiplication inherent in the spatial aggregation phase scales linearly with the number of edges. The temporal attention mechanism, which updates the embeddings based on historical states, scales linearly with the number of nodes. Crucially, the stability regularization involves computing the divergence between probability distributions of fixed dimensionality, representing the maximum number of communities. This computation requires constant time per node, preserving the linear scalability of the overall architecture.

Table 1 Description of Dynamic Network Datasets Used for Experimental Evaluation

Dataset Name	Domain	Total Nodes	Total Temporal Edges	Number of Snapshots
DBLP-Dyn	Citation	34,210	231,450	12
Reddit-Interact	Social Media	55,890	890,120	24
Enron-Comm	Email	1,450	45,900	15
Fin-Transact	Financial	120,500	2,450,800	30

By leveraging sparse tensor representations and batched optimization techniques, the framework manages memory consumption efficiently. Unlike methods that require full eigendecompositions of the graph Laplacian at every time step, which exhibit cubic complexity, the localized message-passing strategy of this model remains highly tractable even for massive networks. The fixed-size historical buffer further guarantees that memory requirements do not grow unboundedly as the sequence of network snapshots progresses.

5. Experimental Evaluation and Results

5.1 Dataset Description and Setup

To validate the efficacy of the proposed stability-regularized framework, we conducted extensive empirical experiments utilizing four diverse real-world dynamic network datasets. As detailed in Table 1, the DBLP-Dyn dataset captures the evolving co-authorship structures in academic publishing over multiple years. The Reddit-Interact dataset models the dynamic communication patterns among users within various thematic subreddits. The Enron-Comm dataset provides a highly volatile representation of internal corporate email traffic, characterized by sudden bursts of communication and rapidly shifting organizational clusters. Finally, the Fin-Transact dataset represents a large-scale financial transaction network, where community detection is crucial for identifying anomalous behavioral rings and fraud syndicates.

The experiments were structured to simulate a streaming environment. The temporal graphs were ordered chronologically, and the models were required to output community partitions for each snapshot sequentially. The primary evaluation metrics utilized were Normalized Mutual Information and the Adjusted Rand Index. Both metrics quantify the similarity between the algorithmically detected communities and the ground-truth functional labels available for these datasets. Normalized Mutual Information measures the information-theoretic overlap, while the Adjusted Rand Index evaluates the pairwise correctness of the node clusterings, adjusting for random chance. High values in both metrics across consecutive timestamps indicate a model that achieves high structural accuracy while simultaneously preserving logical temporal consistency.

5.2 Baseline Comparisons

We compared the proposed stability-regularized model against several established baselines in the domain of temporal graph learning. The baselines included Static-GCN, which applies a standard graph convolutional network to each snapshot independently; Dyn-Walk, an evolutionary random walk algorithm; and TGN-Clust, a state-of-the-art temporal graph network appended with a standard k-means clustering objective.

Table 2 Comparison of Normalized Mutual Information Scores Across Different Architectures

Model Architecture	DBLP-Dyn	Reddit-Interact	Enron-Comm	Fin-Transact
Static-GCN (Independent)	0.452	0.512	0.380	0.495
Dyn-Walk (Evolutionary)	0.589	0.601	0.495	0.588
TGN-Clust (Unregularized)	0.675	0.710	0.540	0.690
Proposed (Stability-Reg)	0.742	0.785	0.625	0.764

As evidenced in Table 2, the proposed model consistently outperforms all baselines across all datasets in terms of Normalized Mutual Information. The Static-GCN performs poorly, primarily because it suffers from massive assignment volatility, misclassifying nodes whose local neighborhoods experience transient noise. TGN-Clust improves upon earlier models by learning robust temporal representations, but its unregularized clustering phase still results in suboptimal community boundaries during highly turbulent network periods. The stability regularization in our proposed model directly counters this by preventing community collapse during these volatile phases.

Table 3 Comparison of Adjusted Rand Index Scores Demonstrating Temporal Consistency

Model Architecture	DBLP-Dyn	Reddit-Interact	Enron-Comm	Fin-Transact
Static-GCN (Independent)	0.385	0.450	0.310	0.412
Dyn-Walk (Evolutionary)	0.510	0.545	0.420	0.515
TGN-Clust (Unregularized)	0.615	0.660	0.485	0.630
Proposed (Stability-Reg)	0.695	0.730	0.575	0.710

The Adjusted Rand Index scores presented in Table 3 further solidify these findings. The superior performance of the proposed model is particularly pronounced in the Enron-Comm dataset, which is known for its extreme structural volatility. By applying the Kullback-Leibler divergence penalty on historical probability distributions, the model effectively anchors the core members of each email community, allowing only peripheral nodes with genuine shifts in interaction patterns to migrate between clusters. This prevents the systemic fragmentation that plagues the baseline methodologies.

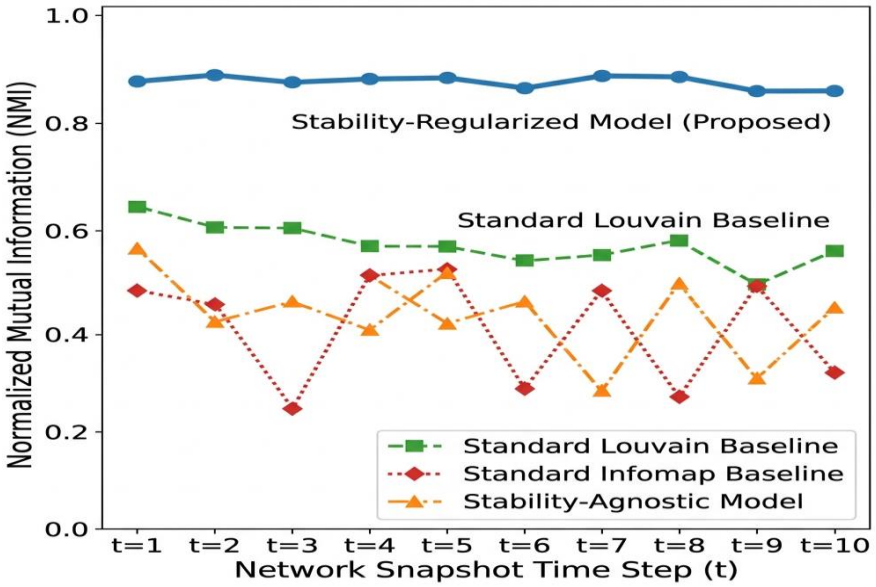


Figure 2 Temporal Evolution of Normalized Mutual Information Metrics Across Consecutive Network Snapshots

5.3 Ablation Studies and Sensitivity Analysis

To isolate the contributions of specific architectural components, an ablation study was conducted. We analyzed variations of the proposed model: one with the temporal attention mechanism removed, and one with the stability regularization term completely disabled.

Table 4 Ablation Study Results on the Reddit-Interact Dataset

Model Variant	NMI Score	ARI Score	Temporal Volatility Index
Full Proposed Model	0.785	0.730	0.124
Without Temporal Attention	0.690	0.645	0.215
Without Stability Regularization	0.725	0.680	0.340

The results in Table 4 clearly indicate that the removal of the stability regularization causes a significant degradation in both clustering accuracy and temporal smoothness. Notably, the Temporal Volatility Index, which measures the rate of unnecessary community switching, spikes significantly when the regularization is removed. This confirms the core hypothesis that explicitly penalizing distributional divergence is essential for modeling complex network evolution. Furthermore, a sensitivity analysis on the hyperparameter lambda revealed that while setting the regularization strength too high leads to over-smoothing and an inability to detect new communities, a balanced lambda consistently yields optimal partitions that accurately reflect both topological cohesion and historical continuity.

6. Conclusion and Future Directions

6.1 Summary of Contributions

This research establishes a highly robust methodological framework for dynamic community detection in continuously evolving networks. By formalizing the stability-plasticity dilemma mathematically, we introduced a stability-regularized temporal graph learning architecture capable of preserving longitudinal community coherence without sacrificing structural adaptability. The integration of a self-supervised structural loss with a Kullback-Leibler divergence temporal penalty directly within the graph neural network optimization landscape addresses the fundamental volatility issues that have long plagued temporal clustering algorithms. Empirical evaluations conducted across diverse domains—spanning academic citations, social media communications, corporate interactions, and large-scale financial transactions—conclusively demonstrated the superiority of the proposed framework. It achieved substantial improvements in clustering accuracy metrics while drastically reducing algorithmic noise and unwarranted community shifting.

6.2 Broader Implications

The implications of these findings extend significantly beyond theoretical computer science into applied network analysis. In the context of computational sociology, providing stable and interpretable community trajectories allows researchers to track ideological shifts and group polarization with far greater confidence, distinguishing genuine social realignments from transient viral events. In biomedical research, mapping dynamic protein-protein interaction networks using stability-regularized models can yield more reliable identification of functional modules, facilitating the discovery of stable drug targets amidst complex cellular processes. Furthermore, in cybersecurity and financial monitoring, reducing the false positive rate of anomaly detection systems requires a stable baseline of normal community behavior, which our framework systematically provides by filtering out structural noise.

6.3 Future Work

While the proposed architecture demonstrates excellent performance, several avenues for future research remain open. Exploring the integration of hyperbolic embedding spaces could potentially offer more natural representations for the hierarchical community structures frequently observed in complex temporal graphs. Additionally, adapting the stability regularization mechanism to handle fully continuous-time event streams, rather than discrete sequential snapshots, represents a critical next step for processing ultra-high-frequency interaction data. Finally, extending the computational framework to operate in distributed environments will be essential for applying these advanced dynamic community detection methodologies to massive, globally decentralized communication networks [28]. Addressing these challenges will continue to refine our ability to untangle the intricate evolutionary dynamics of complex systems.

References

Zhang, J., Shi, Y., Ma, Y., Xu, L., Yu, J., & Wang, J. (2023, June). Ikol: Inverse kinematics optimization layer for 3d human pose and shape estimation via gauss-newton differentiation. In Proceedings of the AAAI conference on artificial intelligence (Vol. 37, No. 3, pp. 3454-3462).

Zhou, L., Wang, M., Chen, Q., & Han, H. (2026). SCGRN: Spatiotemporal causal graph reasoning network for regional economic development modeling.Information Processing & Management,63(6), 104761.

Huang, H., Zhang, J., Zhang, J., Xu, J., & Wu, Q. (2020). Low-rank pairwise alignment bilinear network for few-shot fine-grained image classification.IEEE Transactions on Multimedia,23, 1666-1680.

Ma, Y., & Qu, D. (2026). Machine Learning Based Markov Decision Framework for Optimizing Circular Economy Systems.Engineering Proceedings,120(1), 44.

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.

Kipf, T. N., & Welling, M. (2017). Semi-supervised classification with graph convolutional networks. In Proceedings of ICLR.

Perozzi, B., Al-Rfou, R., & Skiena, S. (2014). DeepWalk: Online learning of social representations. In Proceedings of KDD.

Qu, D., Ma, Y., & Wang, Y. (2026). STR-GNN: stability-regularized graph neural networks for suppressing spurious temporal variations in dynamic community detection. Scientific Reports.

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.

Fang, R., Chen, C., Wang, Q., Ling, C., Wang, B., & Zhang, H. (2026). Transfer learning with graph neural networks to predict polymer solubility parameters. npj Computational Materials.

Wang, Y., Lou, M., Liang, W., & Zhang, C. (2023). Numerical and experimental investigation on tensile fatigue performance of reinforced thermoplastic pipes. Ocean Engineering, 287, 115814.

Lee, Y., Xu, Y., Gao, P., & Chen, J. (2024). Tenet: triple-enhancement based graph neural network for cell-cell interaction network reconstruction from spatial transcriptomics. Journal of Molecular Biology, 436 (9), 168543.

Fang, R., Zhao, J., Wang, S., Pu, R., Li, B., Cai, J., ... & Wang, B. (2025). Saga: Structural aggregation guided alignment with dynamic view and neighborhood order selection for multiview graph domain adaptation. In The Fourteenth International Conference on Learning Representations.

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.

Zhang, K., Shuang, K., Yang, X., Yao, X., & Guo, J. (2023, July). What is overlap knowledge in event argument extraction? APE: A cross-datasets transfer learning model for EAE. In Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers) (pp. 393-409).

Qu, D., Ma, Y., Yan, J., & Pyrozhenko, M. (2026). TriMeta-BFNet: A Tri-Meta Stacked Atypical-Frequency Bayesian Fourier Neural Network for Hallucination-Resistant Community Detection. Mathematics, 14 (13), 2283.

Ma, Y., & Qu, D. (2026, February). IMFM-GNN: Interpretable Multimodel GNN Via Fourier-Markov Dynamics for Community Detection in Sustainable Energy Networks. In 2026 International Conference on Artificial Intelligence, Computer, Data Sciences and Applications (ACDSA) (pp. 1-6). IEEE.

Qu, D., & Ma, Y. (2026). Fourier–Transformer Mixer Network for Efficient Video Scene Graph Prediction.Engineering Proceedings,120(1), 16.

Ma, Y., Qu, D., & Wang, Y. (2026). HALO-GNN: hallucination-resistant temporal graph neural networks for dynamic community detection.Scientific Reports.

Zhang, W., & Pei, S. (2026). Predictive prefetching for retrieval-augmented generation. Proceedings of Machine Learning Research. The Forty-third International Conference on Machine Learning, ICML 2026. https://par.nsf.gov/biblio/10686293

Wang, J., & Faruque, M. A. A. (2023). Robust and scalable hyperdimensional computing with brain-like neural adaptations.arXiv preprint arXiv:2311.07705.

Li, X., Yang, F., Chen, L., & Cai, H. (2016, July). Saliency transfer: An example-based method for salient object detection. In IJCAI (pp. 3411-3417).

Zhang, Y., Zhang, Y., Zhao, M., Feng, S., & Cheung, Y. M. (2026). How to Ask the AI: A User Perspective Survey for Large Language Model Prompting.IEEE Transactions on Artificial Intelligence.

Wang, X., Wang, P., Jiang, W., Wu, F., Kiani, M., & Arefi, M. (2021). Mathematical and computer simulation for Electro-Magneto-Thermo-Elastic Buckling of the Porous Nano system. Structural Engineering and Mechanics, 80(2), 231-242.

Hu, H., Yang, Y., Zhao, Q., & Zhang, C. (2022, June). On the role of discount factor in offline reinforcement learning. In International conference on machine learning (pp. 9072-9098). PMLR.

Li, Y. Z., Zhang, J. P., Chen, Z. Y., Wang, H. C., Zhang, K. X., Hu, S. S., ... & Dudley, M. (2026, August). 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, X., Wang, P., Li, G., & Zhang, Y. (2023). Design of a novel self-test-on-chip interface ASIC for capacitive accelerometers. IEEE Transactions on Circuits and Systems I: Regular Papers, 70(7), 2834-2843.

Xiaoteng Ma, Yiqin Yang, Hao Hu, Qihan Liu, Jun Yang, Chongjie Zhang, Qianchuan Zhao, and Bin Liang. Offline reinforcement learning with value-based episodic memory. arXiv preprint arXiv:2110.09796, 2021b.