AAAI 2026 Supplementary Material

Anonymous submission

Empirical Analysis Details

Graphlet Analysis Algorithm

To uncover the fundamental structural patterns of the RPT network, we counted the occurrences of graphlets in the graph. Graphlets are small, non-isomorphic, induced subgraphs, which have been widely used to capture the local topological structure of nodes in real-world networks (Sarajlić et al. 2016). The frequency distribution of these graphlets provides a structural signature of the network, revealing its underlying organizational principles and distinguishing it from a random graph. The task, formally called subgraph enumeration, is computationally intensive, especially on large graphs. To accelerate the process, recent research has shifted towards utilizing the substantial computational capabilities of GPUs. In our analysis, which focuses on the frequency of connected 4-node graphlets, the state-of-the-art GPU-based subgraph enumeration algorithm EGSM (Sun and Luo 2023) is employed. Specifically, to improve the efficiency of the dominant component—set intersection—during enumeration, EGSM introduces a novel data structure, the Cuckoo Trie, which improves memory coalescing when accessing edge lists. Besides, a hybrid BFS-DFS search strategy is adopted to improve parallelism and ensure scalability for large-scale graphs. By applying this algorithm to our RPT graph snapshots, we obtained a quantitative distribution of these fundamental structural patterns. It was through this analysis that we made a key empirical discovery, as presented in the main text: the star-like graphlet (a central node connected to several peripherals) appeared with a frequency that overwhelmingly dominated all other structures. This data-driven finding is the direct evidence that led us to conclude the prevalence of a core-periphery structure in RPT networks, which in turn motivated our entire role-based modeling approach.

Definition of the Neighborhood Aggregation Factor (NAF)

In the RPT network, we treat entities involved in the transaction process as nodes and construct the graph based on whether there is an RPT between the entities. Central nodes (systemically important financial institutions) typically play a critical role in transaction decision-making and capital flow management, while peripheral nodes (ordinary entities) often serve as information receivers or transaction executors. To empirically examine the various impacts of the nodes, we perform experimental analysis on the functions of central and peripheral nodes during message passing. Specifically, we combine two metrics (1) neighborhood entropy and (2) center-neighbor similarity (Xie et al. 2020) to derive the **neighborhood aggregation factor (NAF)**, which integrates the diversity and similarity of a node's neighborhood. NAF indirectly reflects the information influence during the message passing process.

(1) Neighborhood entropy is a metric measuring the diversity of a node's neighborhood:

$$S_{etp}(v) = -\int_{\mathbb{X}} f^{\mathcal{N}(v)}(x) \cdot log(f^{\mathcal{N}(v)}(x))dx$$

where $\mathbb{X}$ is the feature space, $\mathcal{N}(v)$ is the neighbor set of v, and f is the probability density function.

(2) Center-neighbor similarity is a metric measuring the similarity between a node and its neighbors:

$$S_{sim}(v) = P(\mathcal{N}(v) \mid v) = \frac{1}{|\mathcal{N}(v)|} \sum_{u \in \mathcal{N}(v)} \frac{x_v^\top x_u}{\sum_{k \in V} x_v^\top x_k}$$

where x_v is the node feature of v and V is the node set of the graph.

We combine these two metrics to derive the comprehensive index NAF:

$$NAF(v) = \sigma(1 - Norm(S_{etp}(v))) \cdot \sigma(1 - Norm(S_{sim}(v)))$$

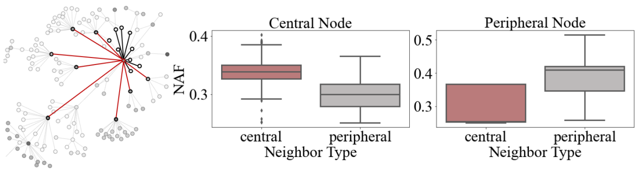
where $Norm(\cdot)$ is a normalization function that scales input values to the range [0, 1], and $\sigma(\cdot)$ is an activation function such as the sigmoid function.

Empirical Analysis of RPT Networks

Prevalence of Core-Periphery Topology. RPT networks are complex relational systems, which implies the structure is not random but governed by underlying generative principles. To uncover the fundamental connection patterns, we performed a graphlet analysis on the RPT dataset by counting the frequencies of all connected 4-node graphlets listed in Figure 1(a). It shows that star-like graphlets, a central node connected to several peripherals, are overwhelmingly dominant. This prevalence of star-like structures is the direct

Graphlet						
Frequency	7,572,014	2,295,757	396,456	17,597	318	0

(a) Occurrences of all 4-node connected graphlets on the RPT network with node roles indicated in different colors.



(b) The layout of a typical RPT network. Node color denotes neighborhood aggregation factor (NAF). The box plot shows the NAF distribution for central and peripheral nodes on RPT dataset.

Figure 1: RPT network exhibits a core-periphery structure, and structural roles count in determining node interactions.

topological signature of a core-periphery business ecosystem, reflecting how a central entity organizes RPT with numerous dependent subsidiaries.

The Implications of Structural Roles. Building on the discovery of core-periphery structure, we further investigate the functional implications of structural roles within the context of GNN message passing. As shown in Figure 1(b), we computed the NAF for each node and analyzed its distribution across structural roles. The results reveal a clear distinction in NAF distributions between central and peripheral nodes, with nodes of the same role exhibiting similar distributions and those of different roles showing significant divergence. Locally, nodes with different roles contribute differently to message propagation, while globally, nodes with similar roles perform comparable functions. (Donnat et al. 2018) reached a similar conclusion by analyzing wavelet coefficients. In addition, structural roles not only influence a node's function but also shape its interactions with neighbors. The message passing process may vary substantially depending on the structural role of neighboring nodes. These observations underscore the importance of incorporating role-based distinctions into models to enhance network analysis and prediction accuracy.

Experiment Details

Our source code is available at GitHub[1].

Datasets

We conducted experiments on four real-world financial transaction networks.

- **RPT:** The dataset was collected from regularly disclosed RPT data in China's nationwide financial market from January 2015 to December 2021. It includes profiles of listed companies and related parties, RPT operation information, and financial exchanges involved in these transactions. In the graph, nodes represent listed companies or related parties, and an edge can be considered as an indication of a related-party transaction between two nodes. This dataset involves over 5,000 listed companies and more than 30,000 enterprises and individuals engaged in the RPT. After data cleaning, we obtained over 670,000 high-quality valid RPT records. We are currently working closely with our partners on data anonymization and open sourcing, aiming to contribute this valuable industry-level dataset to the academic community and promote research on RPT.
- **Elliptic**[2]: The dataset maps bitcoin transactions to real entities belonging to licit categories versus illicit ones. A node in the graph represents a transaction, an edge can be viewed as a flow of Bitcoins between one transaction and the other.
- **Bitcoin OTC**[3]: It is a who-trusts-whom network of people who trade using Bitcoin on a platform called Bitcoin OTC. Nodes represent Bitcoin users, edges represent ratings between users, and it can be used to predict whether a user will rate another user in the next time step.(Kumar et al. 2016, 2018)
- **Bitcoin Alpha**[4]: The dataset was created in the same way as the Bitcoin Alpha dataset, except that the users and ratings come from a different trading platform.

Table 1: Occurrences of all 4-node connected graphlets across datasets with node roles indicated in different colors.

Graphlet						
RPT	7,572,014	2,295,757	396,456	17,597	318	0
Eliiptic	19,798,592	548,224	77,824	480	416	0
OTC	411,544	309,680	123,456	9,460	16,704	152
Alpha	209,800	191,096	77,824	6,410	10,944	112

These financial networks are complex relational systems shaped by human behavior and exhibit core-periphery structure patterns. Specifically, we counted the frequency of 4-node connected graphlets in each dataset in Table 1. The results provide strong empirical evidence: the core-periphery structure is a dominant and general characteristic of financial networks. This pattern is particularly stark in the RPT and Elliptic datasets, where the count of the star-like graphlet is dominantly higher than any other structure. It indicates a clear topology organized around central hub entities. While the star-like graphlet is also most frequent in the Bitcoin OTC and Alpha datasets, the high prevalence of path-like graphlets is also notable. It is important to note that in a financial context, such transaction chains often represent capital flows directed by a core entity, acting as another manifestation of the core-periphery structure rather than a contradiction to it. Nevertheless, the consistent dominance of the

[1]https://anonymous.4open.science/r/RPATGN-79FE

[2]https://www.kaggle.com/datasets/ellipticco/elliptic-data-set

[3]http://snap.stanford.edu/data/soc-sign-bitcoin-otc.html

[4]https://snap.stanford.edu/data/soc-sign-bitcoin-alpha.html

star-like graphlet across diverse financial datasets validates our decision to focus on role-based analysis. It motivates the design of a role perceptual model, which is fundamentally tailored to capture the most prevalent structural pattern in the RPT network.

Baselines

We compare the performance of our proposed method with the state-of-the-art baselines, including two advanced static network embedding models and several temporal graph learning models.

- GAE (Kipf and Welling 2016): Encodes node features using a GCN and reconstructs the graph's adjacency matrix using an inner product decoder.
- VGAE (Kipf and Welling 2016): An extension of GAE that learns the latent representation of nodes through variational inference, better handling the uncertainty in graph data.
- DySAT (Sankar et al. 2020): The model employs attention mechanisms along spatial and temporal dimensions, using GAT and transformers respectively.
- EvolveGCN (Pareja et al. 2020): EvolveGCN adapts the GCN model over the time dimension by evolving GCN parameters through RNNs to capture the dynamics of graph sequences.
- GRUGCN (Seo et al. 2018): The model combines CNNs and RNNs for graph-structured data to simultaneously recognize spatial structures and dynamic patterns.
- VGRNN (Hajiramezanali et al. 2019): It adds higher-order latent variables to the graph recurrent neural network to model the complex dynamics of the graph.
- HTGN (Yang et al. 2021): A node representation learning framework based on hyperbolic geometry, mapping temporal graphs to hyperbolic space to learn temporal regularities, topological dependencies, and implicit hierarchical organization.
- DGCN (Gao et al. 2022): A dynamic graph representation learning model that combines GCN and LSTM to capture both global structure and temporal properties, using a novel Dice similarity measure for node aggregation.
- HGWaveNet (Bai et al. 2023): A hyperbolic graph neural network that leverage the fitness between hyperbolic spaces and data distributions for temporal link prediction.
- HMPTGN (Le and Ta 2024): A dynamic graph learning model that operating directly on the hyperbolic manifold, using manifold-preserving feature transformations and temporal attention.
- RTGCN (Du et al. 2024): The model utilizes global structural role information for temporal graph representation learning, combining structural role-based hypergraphs, GNN, and GRU modules.
- Hawkes-GNN (Qi et al. 2025): The model fuses input snapshots into a single temporal graph and integrates Hawkes processes with GNNs to model temporal edges efficiently.

Settings

We conducted the experiment five times and took the average results. For datasets without node features, we use one-hot encoding as the input feature. We set the embedding dimension as 128 for RPT, Elliptic datasets and 64 for Bitcoin-OTC, Bitcoin-Alpha datasets. We uniformly set the historical snapshot recording window to 5. When determining node roles, we follow the "Pareto principle". We consider the top 20% nodes by link count to be the central roles, while the remaining nodes are peripheral. For model evaluation, we label all known edges in the test snapshots as true, and sample an equal number of non-links as false to balance the classes. During training, we used early stopping with a patience of 50 and a maximum of 200 iterations to prevent overfitting and optimize performance. Our method is implemented using PyTorch 1.12.1 with CUDA 11.3 and Python 3.7 as the backend. The model is trained on a server with four 32GB NVIDIA Tesla V100 GPUs.

Additional Results

Case Study of Company A (2017-2021)

We monitored the dynamic evolution of company A in the RPT networks over time. Figure 2 displays the UMAP projections of the node embeddings around Company A for each quarter from 2017 to 2021. The model was trained on graph snapshots from 2017Q1 to 2019Q4 and evaluated on its ability to predict the network's evolution from 2020Q1 to 2021Q4. The visualizations for 2020-2021 are derived from the node embeddings predicted by our proposed RPATGN for future snapshots.

The visualizations reveal the trajectory of company A within the RPT network. During the initial phase, from 2017 through early 2018, Company A's position was consistently located within the relatively peripheral cluster, suggesting relatively limited RPT activity and fewer links to core entities. Then, its embedding trends towards the central cluster, mirroring the company's real-world strategy of aggressive expansion and the establishment of new RPTs with other core entities. Through training on historical data from 2017 to 2019, our RPATGN learned the distinct pattern of Company A's expansion and role transition. When forecasting the 2020-2021 period, the model accurately predicted that Company A's latent representation would stabilize near the core adjacent neighborhood. Public records confirm that by 2021, company A's expansion and strategic repositioning had stabilized, which is consistent with the model's predictions. By understanding that Company A's function in the network was changing, RPATGN was able to make highly accurate predictions about its new, stable set of RPT links during the test period. This case vividly demonstrates that RPATGN is capable of modeling and anticipating role transitions in dynamic financial networks. By capturing the temporal evolution of network structures and entity roles, the model provides valuable forward-looking insights that can support proactive financial supervision and risk detection.

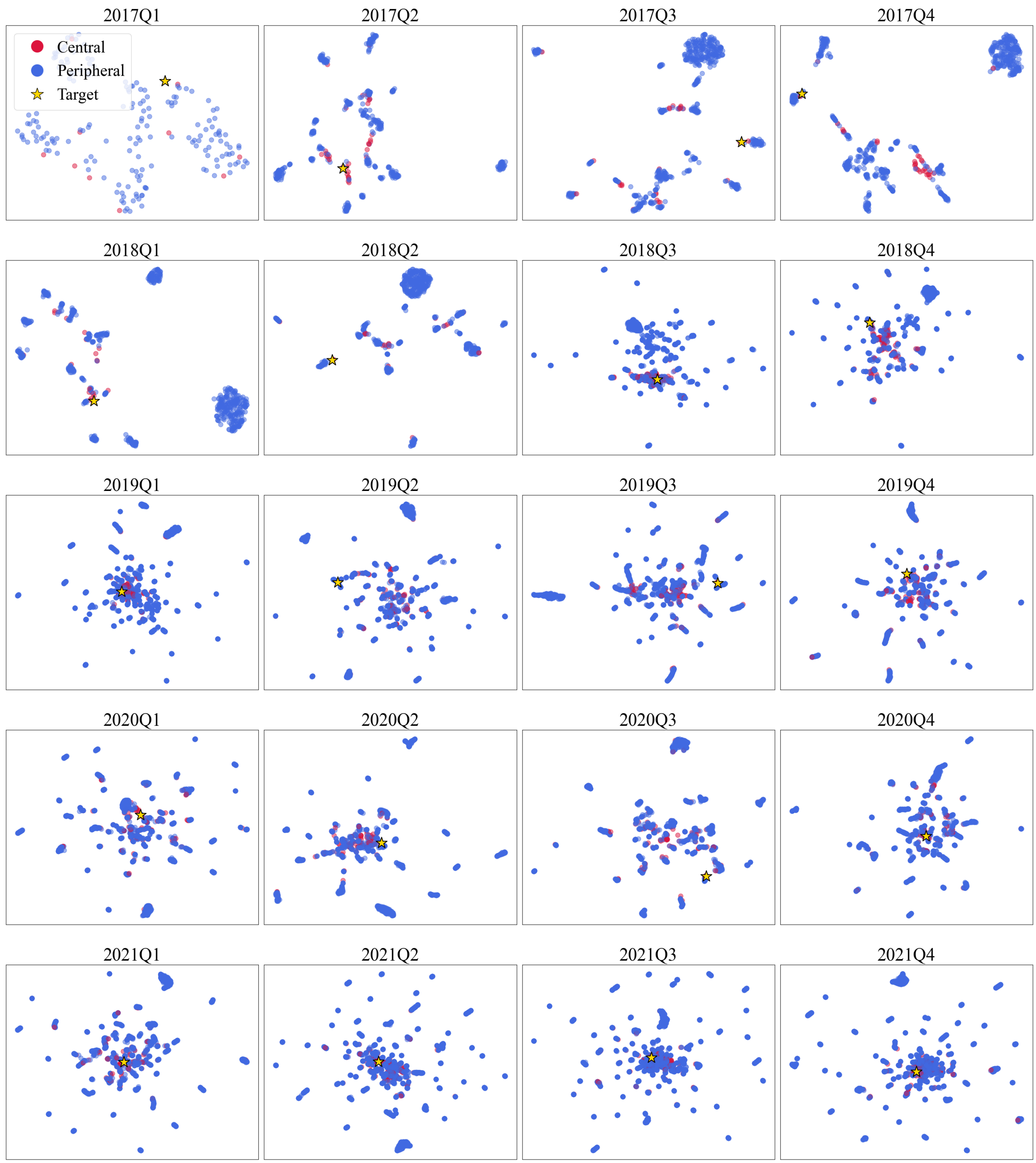


Figure 2: Evolution Trajectory of Company A from 2017 to 2021. This figure presents a sequence of UMAP visualizations for the local subgraph embedding surrounding Company A for every quarter from 2017 to 2021.

References

Bai, Q.; Nie, C.; Zhang, H.; Zhao, D.; and Yuan, X. 2023. Hgwavenet: A hyperbolic graph neural network for temporal link prediction. In *Proceedings of the ACM Web Conference 2023*, 523–532.

Donnat, C.; Zitnik, M.; Hallac, D.; and Leskovec, J. 2018. Learning structural node embeddings via diffusion wavelets. In *Proceedings of the 24th ACM SIGKDD international conference on knowledge discovery & data mining*, 1320–1329.

Du, H.; Shi, L.; Chen, X.; Zhao, Y.; Zhang, H.; Yang, C.; Zhuang, F.; and Kou, G. 2024. Representation Learning of Temporal Graphs with Structural Roles.

Gao, C.; Zhu, J.; Zhang, F.; Wang, Z.; and Li, X. 2022. A novel representation learning for dynamic graphs based on graph convolutional networks. *IEEE Transactions on Cybernetics*, 53(6): 3599–3612.

Hajiramezanali, E.; Hasanzadeh, A.; Narayanan, K.; Duffield, N.; Zhou, M.; and Qian, X. 2019. Variational graph recurrent neural networks. *Advances in neural information processing systems*, 32.

Kipf, T. N.; and Welling, M. 2016. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308*.

Kumar, S.; Hooi, B.; Makhija, D.; Kumar, M.; Faloutsos, C.; and Subrahmanian, V. 2018. Rev2: Fraudulent user prediction in rating platforms. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, 333–341. ACM.

Kumar, S.; Spezzano, F.; Subrahmanian, V.; and Faloutsos, C. 2016. Edge weight prediction in weighted signed networks. In *Data Mining (ICDM), 2016 IEEE 16th International Conference on*, 221–230. IEEE.

Le, V. Q.; and Ta, V. C. 2024. Toward a manifold-preserving temporal graph network in hyperbolic space. In *Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence*, 4380–4388.

Pareja, A.; Domeniconi, G.; Chen, J.; Ma, T.; Suzumura, T.; Kanezashi, H.; Kaler, T.; Schardl, T.; and Leiserson, C. 2020. Evolvegcn: Evolving graph convolutional networks for dynamic graphs. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, 5363–5370.

Qi, Q.; Chen, H.; Cheng, M.; and Liu, H. 2025. Input Snapshots Fusion for Scalable Discrete-Time Dynamic Graph Neural Networks. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, 1138–1149.

Sankar, A.; Wu, Y.; Gou, L.; Zhang, W.; and Yang, H. 2020. Dysat: Deep neural representation learning on dynamic graphs via self-attention networks. In *Proceedings of the 13th international conference on web search and data mining*, 519–527.

Sarajlić, A.; Malod-Dognin, N.; Yaveroğlu, Ö. N.; and Pržulj, N. 2016. Graphlet-based characterization of directed networks. *Scientific reports*, 6(1): 35098.

Seo, Y.; Defferrard, M.; Vandergheynst, P.; and Bresson, X. 2018. Structured sequence modeling with graph convolutional recurrent networks. In *Neural Information Processing: 25th International Conference, ICONIP 2018, Siem Reap, Cambodia, December 13-16, 2018, Proceedings, Part I 25*, 362–373. Springer.

Sun, X.; and Luo, Q. 2023. Efficient gpu-accelerated subgraph matching. *Proceedings of the ACM on Management of Data*, 1(2): 1–26.

Xie, Y.; Li, S.; Yang, C.; Wong, R. C.-W.; and Han, J. 2020. When do gnns work: Understanding and improving neighborhood aggregation. In *IJCAI'20: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence,{IJCAI} 2020*, volume 2020.

Yang, M.; Zhou, M.; Kalander, M.; Huang, Z.; and King, I. 2021. Discrete-time temporal network embedding via implicit hierarchical learning in hyperbolic space. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, 1975–1985.