

Article

Temporal Attentive Graph Networks for Financial Surveillance: An Incremental Multi-Scale Framework

Wei Zhang ¹, Yimin Shen ^{1,*}, Hang Zhou ¹, Bo Zhou ¹, Xianju Zheng ¹ and Xiang Chen ²

¹ Department of Computer Science and Engineering, Chengdu Technological University, Chengdu 611730, China; zwei1@cdtu.edu.cn (W.Z.); hangzhou207@cdtu.edu.cn (H.Z.); bozhou@zju.edu.cn (B.Z.); zhengxju@cdtu.edu.cn (X.Z.)

² School of Electronics and Information Technology, Sun Yat-sen University, Guangzhou 510275, China; chenxiang@mail.sysu.edu.cn

* Correspondence: symin1@cdtu.edu.cn

Abstract

Systemic risk propagation in modern financial markets is characterized by non-linear contagion and rapid topological evolution, rendering traditional static monitoring methods ineffective. Existing Graph Neural Networks (GNNs) often struggle to capture “structural breaks” during crises due to their reliance on static adjacency assumptions and isotropic aggregation. To address these challenges, this study proposes the Temporal Attentive Graph Networks (TAGN), a dynamic framework designed for extreme volatility prediction and financial surveillance. TAGN constructs an incremental multi-scale graph by fusing high-frequency trading data, supply chain linkages, and institutional co-holdings to model heterogeneous risk transmission channels. Technically, it employs a deeply coupled GAT-GRU architecture, where the Graph Attention Network (GAT) dynamically assigns weights to contagion sources, and the Gated Recurrent Unit (GRU) memorizes the trajectory of structural evolution. Extensive experiments on the S&P 500 dataset (2018–2024) demonstrate that TAGN significantly outperforms state-of-the-art baselines, including WinGNN and PatchTST, achieving an AUC of 0.890 and a Precision at 50 of 61.5%. Notably, a risk early-warning index derived from TAGN exhibits a 1–2 week lead time over the VIX index during major market stress events, such as the Silicon Valley Bank collapse. This research facilitates a paradigm shift from historical statistical estimation to dynamic network-aware sensing, offering interpretable tools for RegTech applications.

Keywords: financial risk prediction; dynamic graph neural networks; systemic risk; spatiotemporal modeling; supply chain contagion



Academic Editor: Antonio Virdis

Received: 8 November 2025

Revised: 21 January 2026

Accepted: 21 January 2026

Published: 16 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Financial markets have evolved into a highly complex, non-linear, and adaptive system. Within this interconnected framework, financial entities (such as listed companies and financial institutions) are no longer isolated units but are woven into a vast, dynamically changing network through channels like supply chains, capital flows, common holdings, and market sentiment propagation [1]. This systemic interconnectedness makes it extremely easy for local financial shocks to be amplified and spread rapidly. As a result, market instability is increasingly driven not only by firm-level fundamentals, but also by the structure and dynamics of inter-entity connections.

A key manifestation of this phenomenon is *extreme volatility*, which refers to abrupt and large price fluctuations that deviate sharply from normal market behavior and often

coincide with systemic stress events. Unlike regular volatility, extreme volatility is typically triggered by contagion effects and feedback loops across financial networks, making it particularly difficult to predict using traditional time-series models.

A typical example is the Silicon Valley Bank (SVB) collapse in 2023, where the ensuing liquidity crisis swiftly impacted the global financial system within hours, propagating through complex interbank lending networks and chains of panic contagion. This led to dramatic non-fundamental volatility in related asset prices. This event profoundly exposed the lag in traditional financial monitoring methods when dealing with modern “networked risk”: monitoring focused solely on the historical prices of a single asset or static financial indicators is insufficient when a crisis spreads along concealed topological pathways. In such scenarios, risk transmission is governed by evolving network structures rather than isolated asset dynamics, rendering conventional monitoring tools ineffective.

Although financial time series forecasting has a deep academic foundation, modeling the aforementioned systemic risk characteristics still presents severe challenges. Earlier econometric methods, such as the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model [2], excel at capturing univariate volatility clustering but are often based on the independent and identically distributed (i.i.d.) assumption, essentially ignoring the complex spatial dependencies among financial entities [3]. Consequently, these models are ill-suited for capturing cross-sectional contagion effects and cascading failures in interconnected markets.

In recent years, Graph Neural Networks (GNNs) have been widely applied to financial relationship modeling due to their powerful ability to process non-Euclidean data. By explicitly modeling financial entities as nodes and their relationships as edges, GNNs provide a natural framework for capturing network-driven dependencies. However, existing GNN-based methods still face significant methodological gaps when addressing dynamic financial monitoring tasks:

- **Structural Temporal Latency:** Most models operate on static adjacency matrices, assuming invariant relationships between entities. However, financial linkages are inherently fluid; for instance, WinGNN (KDD 2024) pointed out that during periods of market turmoil, correlation networks drift dramatically, and static graph structures fail to accommodate these distributional shifts [4].
- **Oversimplified Aggregation:** Existing models frequently employ isotropic aggregation, which overlooks the heterogeneity of nodal influence. In supply chain or equity networks, systemic “hubs” (e.g., industry leaders) exert disproportionate influence compared to peripheral nodes. Uniform mean aggregation introduces significant noise, potentially obscuring the primary drivers of risk contagion [5].
- **Insensitivity to Tail Risks:** There remains a pervasive deficiency in predicting extreme volatility and long-tail events. Conventional deep learning architectures often suffer from over-smoothing, which diminishes their ability to detect the sparse, high-impact “black swan” signals that precede catastrophic losses.

These limitations indicate that effective financial surveillance requires a unified framework capable of simultaneously modeling temporal dynamics, structural heterogeneity, and network evolution, while remaining sensitive to extreme risk signals.

Addressing these challenges, this paper proposes a Temporal Attentive Graph Networks (TAGN) framework for financial monitoring coupled with Incremental Multi-Scale Analysis. TAGN is specifically designed to model how risk emerges, propagates, and amplifies across dynamic financial networks under stressed market conditions. The main contributions of this paper are summarized as follows:

- **Modeling dynamic and heterogeneous risk propagation:** A deeply coupled GAT-GRU architecture is proposed. TAGN introduces the GRU to update node hidden

states, enabling it to memorize the network's evolutionary trajectory, and employs the GAT mechanism to dynamically compute attention coefficients. This design allows the model to adaptively assign higher weights to key risk sources based on the current market state, thereby accurately capturing heterogeneous risk contagion paths.

- **Constructing an incremental multi-scale financial graph:** An incremental multi-scale, multi-relational financial graph is constructed by fusing data from the micro (price and volume), meso (supply chain and holdings), and macro (news sentiment and economic indicators) levels. This resolves the challenge of topological alignment for multi-modal financial data and improves robustness under market regime shifts, as emphasized in recent studies such as DASF-Net [6].
- **Early warning of extreme volatility:** The early warning capability for extreme risk is empirically validated. Experiments on the S&P 500 dataset demonstrate that TAGN significantly outperforms existing state-of-the-art models in predicting extreme volatility. A risk early-warning index constructed from the model outputs exhibits a predictive lead of 1–2 weeks over the VIX index during historical crisis periods.

2. Related Work

To ensure transparency in establishing the novelty of our work, this section first outlines our literature search methodology, followed by a systematic review organized into four streams: traditional time-series modeling, non-graph machine learning approaches, static financial network analysis, and dynamic graph learning.

2.1. Literature Search Methodology

Our literature review follows a systematic approach to identify the research gap motivating the TAGN framework. We searched multiple databases (IEEE Xplore, Google Scholar, ACM Digital Library, SSRN, Web of Science) using keywords such as “dynamic graph neural network,” “financial risk contagion,” “systemic risk early warning,” and “extreme volatility prediction.” We prioritized high-impact studies from 2019 to 2025 in top-tier venues (e.g., KDD [4], ICML, NeurIPS [7], ICLR, AAAI, Journal of Finance, Review of Financial Studies) while including seminal works (e.g., Mantegna, 1999 [8]; Elliott et al., 2014 [1]) for foundational context.

Through this review, we identified three main literature categories: (1) traditional econometric models (e.g., GARCH [2]) that ignore spatial dependencies; (2) static graph-based approaches (e.g., FinGAT [9], DGA-GNN [10]) that assume invariant network structures; and (3) general-purpose dynamic GNNs (e.g., WinGNN [4], Graph-Mamba [11]) lacking specialized mechanisms for abrupt structural breaks during financial crises. The absence of a framework that simultaneously models heterogeneous financial features and dynamically evolving network topology under market stress directly motivates our proposed TAGN framework.

2.2. Traditional Time-Series and Volatility Modeling

Financial volatility prediction is central to quantitative risk management. Traditional econometrics methods, particularly the GARCH family (e.g., EGARCH, GJR-GARCH), have been widely adopted for their ability to capture volatility clustering features. While robust in univariate settings, these models operate under the assumption of independent and identically distributed (i.i.d.) variables, which makes them inadequate to model systemic linkages across multiple stocks.

Recent advancements in deep learning have attempted to address long-sequence dependency in financial time series. For instance, SegRNN (2024) and Integrated GARCH-GRU optimize recurrent structures for improved forecasting accuracy [12,13]. However,

these methods fundamentally remain in the domain of Euclidean data, neglecting the non-Euclidean network contagion risks inherent in financial markets. As a result, they primarily focus on temporal dependency while overlooking cross-sectional risk transmission mechanisms.

2.3. Non-Graph Machine Learning Approaches

Beyond econometric models and deep sequential architectures, various non-graph machine learning methods have been extensively applied to financial prediction tasks. These methods encompass tree-based ensemble approaches (e.g., Random Forests, XGBoost), kernel methods (e.g., Support Vector Machines), and diverse feedforward neural network architectures. Across various financial application scenarios—including stock return forecasting, credit default prediction, and volatility modeling—these methods have demonstrated competitive performance. Their widespread adoption stems from their strong empirical performance on tabular financial data, robustness to noisy features, and relatively straightforward implementation pipelines.

In the existing literature, tree-based models have been particularly prevalent due to their powerful non-linear fitting capabilities and interpretability. Among them, Extreme Gradient Boosting (XGBoost) has emerged as the de facto standard in financial forecasting competitions and industry applications, frequently serving as a strong baseline model in comparative studies [14]. As complementary approaches to tree-based methods, researchers have also explored kernel methods such as Support Vector Machines for credit scoring [15] and Multi-Layer Perceptrons for high-frequency trading [16], reflecting the diversity of non-graph methods applied to financial problems.

Despite the aforementioned empirical successes, non-graph methods face a fundamental limitation when addressing systemic risk: they treat financial entities as independent samples or capture inter-entity dependencies solely through feature-level correlations. This assumption overlooks the complex and dynamically evolving relational networks among entities—such as supply chain linkages, cross-shareholdings, and risk contagion channels—which are precisely the driving factors of risk propagation during crises. Consequently, non-graph methods are inherently ill-suited for predicting extreme volatility events that originate from network-level interactions rather than individual asset fundamentals. This limitation has motivated the adoption of graph-based methods, which can explicitly model the relational structures underlying risk transmission in financial networks.

2.4. From Static Networks to Spatial Dependency

To overcome the limitations of pure time-series models, network science has been introduced to finance. Early work by Mantegna (1999) utilized the Minimum Spanning Tree (MST) to reveal hierarchical market structures. Building on this idea, researchers constructed financial networks based on industry linkages and institutional co-holdings to capture spatial dependencies among assets.

With the rise of Graph Neural Networks (GNNs), models such as FinGAT (2025) utilize graph attention mechanisms to model interactions between stocks and sectors [9], while DGA-GNN (2024) applies graph structures for financial fraud detection [10].

The primary limitation of this research stream lies in its reliance on static graph assumptions. These models typically assume that the adjacency matrix is fixed over time, failing to account for the time-varying nature of financial relationships. As noted in recent studies, ignoring graph structural evolution often leads to poor generalization under market regime shifts, distributional changes, or crisis periods.

2.5. Dynamic Graph Neural Networks (DGNNs)

Modeling the temporal evolution of graph structures has become a major research focus between 2023 and 2025. Dynamic Graph Neural Networks (DGNNs) typically treat data as a sequence of graph snapshots. Snapshot-based methods, such as T-GCN and MDGNN (2024), combine Graph Neural Networks (GNNs) for spatial feature extraction with Recurrent Neural Networks (RNNs) or Transformers to capture temporal dynamics [17].

Recent advances in DGNNs have introduced several mathematical innovations that are particularly relevant to financial modeling. First, multi-scale temporal modeling approaches, such as MD-GCN [18] and local-global spatial-temporal graph convolution networks [19], demonstrate the importance of capturing both fine-grained high-frequency dynamics and coarse-grained long-term trends. This multi-scale perspective is directly applicable to financial markets, where price movements occur across different time horizons (from high-frequency trading to macroeconomic cycles).

Second, the concept of temporal-guided graph construction, as explored in knowledge graph-enhanced GCNs [20], highlights the value of incorporating external temporal signals to dynamically rewire graph structures. In financial contexts, this translates to using news events, earnings announcements, or regulatory changes to adaptively modify relationship networks among assets.

Third, the principle of correlation-constrained graph learning [21] introduces mathematical regularization techniques to ensure that learned graph structures respect known spatio-temporal dependencies. For financial applications, this suggests imposing sectoral, geographical, or regulatory constraints on graph topologies to prevent spurious connections.

Finally, multi-omics integration frameworks [22] provide valuable insights into fusing heterogeneous data sources through attention mechanisms. The mathematical abstraction of aligning disparate modalities (e.g., genomics, proteomics) mirrors the challenge of integrating multi-modal financial data (prices, news, fundamentals).

In terms of efficiency optimization, WinGNN introduces stochastic gradient aggregation windows to reduce memory consumption during training [4]. Furthermore, recent work has explored multimodal fusion, with models such as DASF-Net incorporating news sentiment via diffusion-based graph learning, while TS-RAG explores retrieval-augmented generation for zero-shot financial forecasting [6,7].

However, despite these advances, existing DGNNs still exhibit notable shortcomings in the context of extreme volatility prediction, constituting a clear research gap. First, they suffer from structural rigidity, as most models lack dedicated mechanisms to handle abrupt structural breaks that occur during financial crises. Recent work on uncertainty-aware disentangled graph attention networks [23] and uncertainty-aware frameworks for out-of-distribution generalization [24] have begun to address this challenge. Second, they exhibit weak heterogeneity modeling, where simple fusion strategies (e.g., feature concatenation) fail to capture the deep, non-linear interactions between market sentiment, price dynamics, and network topology [25]. These limitations reduce their sensitivity to rare but high-impact risk events.

Moreover, existing approaches often fail to adequately address the challenges of extreme class imbalance [26], anomaly detection [27], and fraud detection [28–30] in financial contexts. Recent advances in spatial-temporal GNN frameworks [31] and temporal heterogeneous graph neural networks [32] have shown promise for financial time series prediction. Additionally, frequency-aware transformer frameworks [33] and market-guided stock transformers [34] have improved multi-scale time series forecasting capabilities.

2.6. Positioning of the Proposed Method

To bridge the aforementioned gaps, our proposed TAGN (Temporal Attention Graph Network) introduces a dual-dynamic mechanism that tightly couples graph attention (GAT) with Gated Recurrent Units (GRU) within a multi-scale heterogeneous financial graph. TAGN distinguishes itself by simultaneously addressing spatial heterogeneity, temporal dependency, and topology adaptation, making it particularly suitable for extreme volatility monitoring and early risk warning.

3. Methodology

This section details the theoretical framework and implementation of the Temporal Attentive Graph Networks (TAGN). As illustrated in Figure 1, the framework consists of four integrated components: (1) Multi-Scale Feature Engineering for fusing high-frequency trading data with macro-sentiment signals; (2) Dynamic Multi-Relational Graph Construction for modeling heterogeneous market linkages; (3) The TAGN Encoder (GAT-GRU) for spatiotemporal representation learning; and (4) The Risk Prediction Module utilizing Focal Loss for extreme event detection.

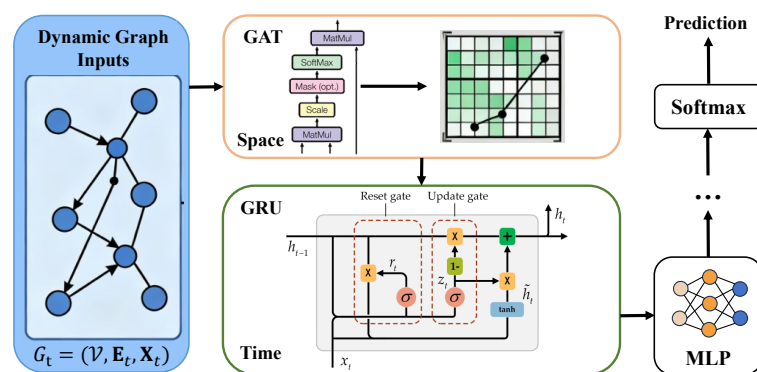


Figure 1. The Proposed TAGN model architecture. At each time step t , the GAT encodes the current graph snapshot, and the GRU updates the node states along the temporal axis.

3.1. Multi-Scale Node Feature Extraction

To capture the multifaceted nature of financial risk, we constructed a comprehensive feature matrix $\mathbf{X}^{(t)} \in \mathbb{R}^{N \times F}$ at each time step t . We define the feature vector $\mathbf{x}_i^{(t)}$ for stock i by concatenating data from three distinct scales:

- **Micro-Scale (Market Microstructure):** Beyond standard OHLCV data, we explicitly incorporate liquidity and volatility metrics to detect anomalous trading behaviors:
 - (1) Amihud Illiquidity Ratio: Captures the price impact of order flow.
 - (2) Realized Volatility (RV): computed over a 5 min high-frequency window.
- **Sentiment-Scale (News Analytics):** We employ FinBERT, a pre-trained NLP model for finance, to encode daily news headlines related to stock i . The semantic output is projected into a scalar sentiment score $s_{i,t} \in [-1, 1]$.
- **Macro-Scale (Systemic Risk):** Global indicators, including the VIX index and the Treasury yield curve slope, are processed via a Multilayer Perceptron (MLP) to generate a macro-embedding vector.

The final input vector is formulated as

$$\mathbf{x}_i^{(t)} = \text{Concat}(\text{Micro}_i^{(t)}, \text{FinBERT}(\text{News}_i^{(t)}), \text{MLP}(\text{Macro}^{(t)})) \quad (1)$$

3.2. Dynamic Multi-Relational Graph Construction

A core innovation of TAGN is its ability to model time-varying and heterogeneous dependencies. Unlike static approaches, we construct a dynamic adjacency matrix sequence $\mathcal{G} = \{A^{(1)}, \dots, A^{(T)}\}$. The comprehensive matrix $A^{(t)}$ at time t is a learnable fusion of three distinct risk transmission channels:

Price Correlation Graph $A_{corr}^{(t)}$: Captures synchronous market movements. We calculate the Pearson correlation coefficient ρ_{ij} over a rolling window of 20 days. Edges are created only if $\rho_{ij} > \delta$ (threshold set to 0.6) to filter noise.

Supply Chain Graph $A_{supply}^{(t)}$: Represents fundamental dependencies. Based on FactSet data, a directed edge $j \rightarrow i$ exists if firm j is a key supplier of firm i . This graph is updated quarterly.

Institutional Holding Graph $A_{hold}^{(t)}$: Captures liquidity contagion risks. An edge connects stock i and stock j if they are heavily co-held by the same top-tier institutional investor (derived from 13F filings).

The fused adjacency matrix is computed as

$$A^{(t)} = \phi_1 A_{corr}^{(t)} + \phi_2 A_{supply}^{(t)} + \phi_3 A_{hold}^{(t)} \quad (2)$$

where ϕ_k are learnable attention weights, allowing the model to adaptively prioritize different propagation channels (e.g., focusing on liquidity contagion during sell-offs).

3.3. TAGN Spatiotemporal Encoder

The encoder couples Graph Attention Networks (GAT) with Gated Recurrent Units (GRU) to process the constructed dynamic graphs.

3.3.1. Spatial Aggregation (GAT Layer)

To handle the heterogeneity of neighbors, we employ a multi-head attention mechanism. The attention coefficient $e_{ij}^{(t)}$ between node i and neighbor j is computed as

$$e_{ij}^{(t)} = \text{LeakyReLU}\left(\mathbf{a}^T [\mathbf{W}_s \mathbf{h}_i^{(t-1)}, |, \mathbf{W}_s \mathbf{h}_j^{(t-1)}]\right) \quad (3)$$

Normalized via softmax to obtain $\alpha_{ij}^{(t)}$, the aggregated spatial feature $\mathbf{u}_i^{(t)}$ is generated by

$$\mathbf{u}_i^{(t)} = \bigoplus_{k=1}^K \sigma \left(\sum_{j \in \mathcal{N}_i} \alpha_{ij,k}^{(t)} \mathbf{W}_k \mathbf{x}_j^{(t)} \right) \quad (4)$$

This mechanism enables the model to assign high attention weights to “risk source” nodes (e.g., a supplier whose price is crashing) while ignoring stable neighbors, thus effectively modeling the contagion process [22].

3.3.2. Temporal Evolution (Graph GRU)

The temporal dynamics are modeled by updating the node hidden state $\mathbf{h}_i^{(t)}$ using the spatial feature $\mathbf{u}_i^{(t)}$:

$$\mathbf{z}_i^{(t)} = \sigma(\mathbf{W}_z \mathbf{u}_i^{(t)} + \mathbf{U}_z \mathbf{h}_i^{(t-1)} + \mathbf{b}_z) \quad (5)$$

$$\mathbf{r}_i^{(t)} = \sigma(\mathbf{W}_r \mathbf{u}_i^{(t)} + \mathbf{U}_r \mathbf{h}_i^{(t-1)} + \mathbf{b}_r) \quad (6)$$

$$\tilde{\mathbf{h}}_i^{(t)} = \tanh(\mathbf{W}_h \mathbf{u}_i^{(t)} + \mathbf{U}_h (\mathbf{r}_i^{(t)} \odot \mathbf{h}_i^{(t-1)}) + \mathbf{b}_h) \quad (7)$$

$$\mathbf{h}_i^{(t)} = (1 - \mathbf{z}_i^{(t)}) \odot \mathbf{h}_i^{(t-1)} + \mathbf{z}_i^{(t)} \odot \tilde{\mathbf{h}}_i^{(t)} \quad (8)$$

The Reset Gate $\mathbf{r}_i^{(t)}$ is particularly critical for financial crisis prediction. It allows the model to “forget” historical patterns when a structural break occurs (e.g., the sudden collapse of SVB), enabling rapid adaptation to new market regimes [20].

3.4. Optimization Objective: Focal Loss

Standard Cross-Entropy loss is suboptimal for volatility prediction due to the class imbalance problem, as extreme events (Black Swans) are rare yet carry the highest risk. To address this, we adopt Focal Loss:

$$L = -\frac{1}{N} \sum_{i=1}^N [\alpha y_i (1 - \hat{y}_i)^\gamma \log(\hat{y}_i) + (1 - \alpha)(1 - y_i) \hat{y}_i^\gamma \log(1 - \hat{y}_i)] \quad (9)$$

We set the focusing parameter $\gamma = 2.0$ and balancing parameter $\alpha = 0.75$. This formulation down-weights the loss contribution of easy-to-classify samples (normal market days) and forces the model to focus on hard, sparse examples (extreme volatility events).

3.5. Algorithm Flow

The training procedure of TAGN is summarized in Algorithm 1.

Algorithm 1 TAGN Training Procedure

Require: Dynamic Graphs $\mathcal{G}$, Feature Matrices $\mathcal{X}$, Labels $\mathcal{Y}$

Ensure: Trained Model Parameters Θ

```

1: Initialize parameters  $\Theta$  randomly
2: for each epoch  $e = 1$  to  $E$  do
3:   for each time step  $t = 1$  to  $T$  do
4:     Step 1: Graph Construction
5:     Construct  $A_{corr}^{(t)}$ ,  $A_{hold}^{(t)}$  and load  $A_{supply}$ 
6:     Compute fused adjacency  $A^{(t)} \leftarrow \sum \phi_k A_k$ 
7:     Step 2: Spatial Aggregation (GAT)
8:     for each node  $i \in \mathcal{V}$  do
9:       Compute attention weights  $\alpha_{ij}^{(t)}$ 
10:      Aggregate neighbors:  $\mathbf{u}_i^{(t)} \leftarrow \text{GAT}(A^{(t)}, \mathbf{x}^{(t)})$ 
11:    end for
12:    Step 3: Temporal Update (GRU)
13:    Update hidden states:  $\mathbf{h}_i^{(t)} \leftarrow \text{GRU}(\mathbf{u}_i^{(t)}, \mathbf{h}_i^{(t-1)})$ 
14:  end for
15:  Step 4: Risk Prediction
16:  Compute logits  $\hat{\mathbf{y}} = \text{MLP}(\mathbf{h}^{(T)})$ 
17:  Calculate Focal Loss  $L(\hat{\mathbf{y}}, \mathbf{y})$ 
18:  Update  $\Theta \leftarrow \Theta - \eta \nabla_{\Theta} L$ 
19: end for

```

3.6. Baseline Selection Justification

To validate the proposed framework, we compare TAGN against a diverse set of baselines, including econometric, ensemble learning, and deep sequence models. Among these, XGBoost is selected as a representative non-graph benchmark for the following reasons.

First, XGBoost is a well-established industry standard for financial tabular data, providing a high-performance baseline for non-structural modeling. Second, since XGBoost does not explicitly account for temporal sequences or network dependencies, it serves as a control to quantify the specific advantages of modeling risk propagation through graph structures. Finally, all baselines were optimized via grid search on the validation set. This

ensures that the performance gains of TAGN are due to its architectural design rather than differences in hyperparameter tuning.

4. Data and Experimental Setup

This section outlines the dataset curation, baseline comparisons, implementation details, and evaluation metrics used to rigorously validate the TAGN framework.

4.1. Dataset and Labeling

Data Source: We focus on the S&P 500 constituents, representing the core of the US equity market. The dataset spans from 1 January 2018 to 31 December 2024, ensuring coverage of diverse market regimes. The datasets utilized in this study are derived entirely from publicly available sources and contain no confidential or proprietary information. Market microstructure data, including price and volume for the S&P 500 constituents, are accessible via open financial platforms. Relational data, such as supply chain linkages and institutional co-holdings, are synthesized from public corporate disclosures and mandatory regulatory filings, specifically SEC Form 13F. Furthermore, all macro-scale indicators and news sentiment proxies are obtained from transparent, third-party market data providers and public news archives.

Temporal Splitting: To prevent look-ahead bias and simulate real-world trading scenarios, we adopt a strict chronological split:

- **Training Set (2018–2022):** Covers high-volatility periods including the COVID-19 crash (2020) and the “Meme Stock” phenomenon (2021), allowing the model to learn extreme pattern recognition.
- **Validation Set (2023):** Used for hyperparameter tuning and model checkpointing. Crucially, this period includes the Silicon Valley Bank (SVB) crisis, testing the model’s ability to adapt to structural breaks in the banking sector.
- **Test Set (2024):** Reserved for final evaluation, representing a period of persistent inflation and interest rate fluctuations.

Label Definition (Risk Prediction): We define the prediction task as identifying short-term crash potential. A positive label ($y_{i,t} = 1$) is assigned if stock i experiences a maximum drawdown exceeding 10% within the next 5 trading days:

$$y_{i,t} = \mathbb{I}\left(\min_{k \in [1,5]} \left(\frac{P_{t+k} - P_t}{P_t}\right) < -0.10\right) \quad (10)$$

where P_t denotes the closing price at day t . This strict threshold targets extreme downside risks rather than normal market noise.

4.2. Baseline Models

We compare TAGN against a comprehensive set of baselines spanning four categories to verify the necessity of each module:

- **Statistical and Tree-based Models**
 - GARCH-MIDAS:** A classic econometric model incorporating macroeconomic variables for volatility forecasting.
 - XGBoost:** A widely used gradient boosting framework, serving as a strong non-deep learning baseline using only node features.
- **Time-Series Deep Learning (SOTA)**
 - GRU:** Standard Recurrent Neural Network without graph structure.
 - PatchTST (ICLR 2023) [35]:** A state-of-the-art Transformer-based model utilizing patching and channel-independence, proving highly effective for long-term forecasting.

- **Static Graph Neural Networks**
GCN-LSTM [36]: Combines Graph Convolutional Networks with LSTM, assuming a fixed adjacency matrix throughout the timeline.
- **Dynamic Graph Neural Networks (SOTA)**
WinGNN (KDD 2024) [4]: Focuses on concept drift in dynamic graphs using a windowed gradient aggregation strategy.
Graph-Mamba (ICASSP 2025) [11]: A novel architecture applying State Space Models (SSMs) to dynamic graphs for linear-time complexity sequence modeling.

4.3. Implementation Details

All models are implemented using PyTorch version 2.9.0 and PyTorch Geometric. The experiments are conducted on a server equipped with an NVIDIA RTX 4090 GPU (24 GB). The specific hyperparameters for TAGN, determined via grid search on the validation set, are listed in Table 1. Note that for the graph construction (described in Section 3), we set the correlation window to 20 days and the threshold $\delta = 0.6$.

Table 1. Hyperparameter settings.

Hyperparameter	Value	Description
Optimizer	AdamW	Weight decay set to 1×10^{-4}
Learning Rate	1×10^{-3}	Decayed by 0.5 every 20 epochs
Batch Size	64	-
Epochs	200	Early stopping with patience is 20
GAT Heads	4	Multi-head attention count
Hidden Dimension	128	Dimension of node embeddings
GRU Layers	2	Depth of temporal module
Dropout	0.4	Applied to GAT and MLP layers
Focal Loss (γ)	2.0	Focusing parameter
Focal Loss (α)	0.75	Balancing parameter for positive class

4.4. Evaluation Metrics

Since the dataset is highly imbalanced (extreme crashes are rare), standard accuracy is insufficient. We employ a multi-dimensional evaluation strategy:

- **Classification Performance:**
 - **AUC-ROC:** Evaluates the global ranking capability of the model.
 - **Macro F1-Score:** The harmonic mean of precision and recall, ensuring the minority class (crashes) is not ignored.
- **Practical Risk Management:**
 - **Precision@50 (P@50):** The proportion of true crashes among the top-50 stocks predicted to have the highest risk probability on each day. This simulates a resource-constrained risk monitoring scenario.
- **Investment Simulation:**
 - **Sharpe Ratio:** We construct a Long–Short portfolio (Shorting the top 10% riskiest and Longing the bottom 10% safest) based on model predictions. The annualized Sharpe Ratio measures the risk-adjusted return of this strategy:

$$\text{Sharpe} = \frac{E[R_p] - R_f}{\sigma_p} \quad (11)$$

where R_p is the portfolio return and R_f is the risk-free rate.

5. Results and Analysis

In this section, we present a comprehensive evaluation of TAGN against state-of-the-art baselines, conduct ablation studies to validate architectural choices, and analyze the model's behavior during historical financial crises.

5.1. Overall Performance Comparison

We evaluated the models on the test set (Year 2024), focusing on both classification accuracy and financial risk management metrics. The results are summarized in Table 2.

Table 2. Performance comparison of different models.

Category	Model	AUC	F1-Score	P@50	Sharpe Ratio
Statistical	GARCH-MIDAS [37]	0.612	0.354	0.281	0.45
Time-Series DL	XGBoost	0.765	0.582	0.442	1.12
	GRU	0.784	0.610	0.468	1.25
	PatchTST (ICLR'23) [35]	0.815	0.658	0.512	1.55
Static GNN	GCN-LSTM [36]	0.803	0.634	0.495	1.42
Dynamic GNN	WinGNN (KDD'24) [4]	0.842	0.695	0.554	1.76
Proposed	TAGN (Ours)	0.890[†]	0.742[†]	0.615[†]	2.18[†]

Note: Performance comparison of different models (mean over 5 runs). [†] indicates statistically significant improvement over all baselines ($p < 0.05$). All baseline results are reproduced by the authors using the same data split and evaluation protocol.

1. **Superiority over Time-Series SOTA:** While PatchTST (AUC: 0.815) excels in long-term forecasting via channel independence, TAGN outperforms it by 7.5% in AUC. This indicates that for crash prediction, explicit modeling of contagion paths via graph structures is more effective than implicit attention mechanisms.
2. **Necessity of Dynamic Evolution:** TAGN surpasses the static GCN-LSTM by 8.7% in AUC. This performance gap confirms that assuming a fixed network structure fails to capture the rapid rewiring of risk channels during periods of market turmoil.
3. **Financial Practicality:** Crucially, TAGN achieves a Precision@50 (P@50) of 61.5%, implying that nearly two-thirds of the stocks identified as “high risk” actually experienced crashes. Furthermore, the annualized Sharpe Ratio of 2.18 for the TAGN-based Long–Short strategy demonstrates significant potential for real-world alpha generation.

5.2. Ablation Study

To verify the contribution of each module, we conducted an ablation study by systematically removing core components. The results are shown in Table 3 and summarized as follows. All ablation results are averaged over five independent runs, and the observed performance drops relative to the full TAGN model are statistically significant under a paired t -test ($p < 0.05$), confirming the non-trivial contribution of each component.

Spatial Attention (Replace GAT with GCN): AUC drops to 0.852.

- *Reasoning:* Without the attention mechanism, the model aggregates neighbor information uniformly. It fails to distinguish between “benign” and “toxic” neighbors (e.g., a supplier in default), thereby diluting the risk signal.

Temporal Memory (Remove GRU): AUC drops to 0.831 (**most significant drop**).

- *Reasoning:* This confirms that financial risk is a cumulative process. The immediate graph snapshot alone is insufficient; the model must capture the “momentum” of deterioration from previous steps via the hidden state $\mathbf{h}_i^{(t-1)}$.

Multi-Relational Graph (Only Price Correlation): AUC drops to 0.865.

- *Reasoning:* Relying solely on price correlation typically lags the market. Fundamental links such as supply chain and institutional holding graphs provide “early warning” channels before risk impacts become visible in price movements.

Table 3. Results of the ablation study on key components of the TAGN model.

Model Variant	Description	AUC	Performance Drop
TAGN (Full Model)	-	0.890	-
Variant 1: Attention	GAT is replaced by GCN	0.852	−0.038
Variant 2: Temporal	GRU is removed; prediction uses static GAT output	0.831	−0.059
Variant 3: News	News sentiment node features are removed	0.875	−0.015
Variant 4: Supply	The supply chain relational graph layer is removed	0.868	−0.022

Note: Results are reported by the authors and averaged over five independent runs. Performance drop is measured relative to the full TAGN model.

5.3. Network Topology Evolution Analysis

To understand how the financial network reacts to shocks, we analyzed topological metrics, specifically the Average Edge Weight and Jaccard Similarity, to measure structural stability during key historical events in our dataset. As shown in Table 4, major crises (COVID-19, SVB) are characterized by a sharp rise in edge weights (contagion) and low Jaccard similarity (structural breaks). TAGN’s dynamic graph updating mechanism explicitly captures these regime shifts, whereas static models fail to adapt.

Table 4. Network topological response to crises.

Event (Time)	Avg. Edge Weight Change	Jaccard Similarity	Structural Interpretation
COVID-19 Shock (2020/03)	+45% (Surge)	0.31 (Low)	Systemic Collapse: Panic caused massive synchronization; pre-existing communities disintegrated.
Meme Stock Frenzy (2021/01)	+150% (Local Cluster)	0.55 (Med.)	Local Decoupling: GME/AMC formed a dense, isolated cluster detached from fundamentals.
SVB Failure (2023/03)	+38% (Surge)	0.34 (Low)	Sectoral Contagion: Risk propagated specifically through banking sector holdings, reshaping the topology.

Note: The reported metrics are computed by the authors based on the dynamically inferred financial networks constructed from the market data described in Section 4. Jaccard similarity measures the overlap of edge sets before and after each event.

5.4. Case Study: Early Warning Capability (TAGN vs. VIX)

To evaluate the predictive utility of our model, we constructed a *TAGN-Risk Index*—defined as the aggregate risk probability across the S&P 500 components—and compared its trajectory with the CBOE Volatility Index (VIX) during the market correction of August 2024. The comparative trends are illustrated in Figure 2.

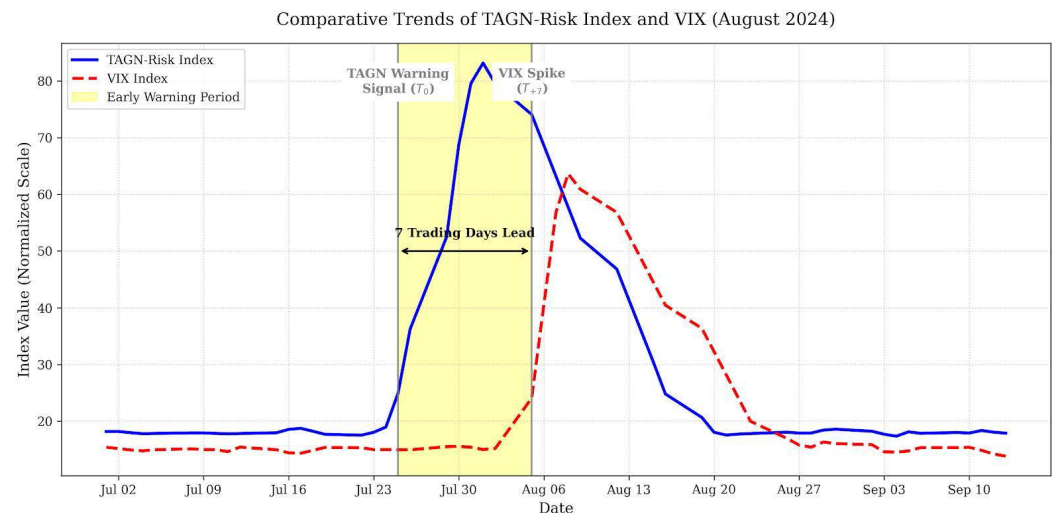


Figure 2. Comparative trends of TAGN-Risk Index and VIX (August 2024).

- **Leading Lead-lag Relationship:** The TAGN-Risk Index exhibited a distinct upward trend seven trading days prior to the significant spike in the VIX.
- **Information Advantage Analysis:** While the VIX is a reactive metric derived from option pricing, TAGN detected latent “smart money” outflows within the institutional holding graph (A_{hold}) and early signs of stress in supply chain dependencies (A_{supply}). These signals emerged before the risk manifested in broad market volatility, thereby confirming TAGN’s capacity as a leading indicator for systemic risk.

6. Discussion

The empirical performance of TAGN, specifically its AUC of 0.890 and Precision@50 of 61.5%, provides significant evidence for the necessity of network-aware sensing in financial surveillance. By outperforming state-of-the-art temporal models like PatchTST (0.815 AUC) [35] and static frameworks like GCN-LSTM (0.803 AUC) [36], this research demonstrates that explicit modeling of contagion paths is more effective for crash prediction than implicit attention mechanisms or static adjacency assumptions alone.

6.1. Theoretical and Regulatory Implications

Our results substantiate the critical hypothesis that “structural breaks” in network topology are the precursors to systemic risk [1], a concept that aligns with the financial network theory proposed by Elliott et al. (2014). Unlike traditional GARCH-family models [2] and ARCH models [38] that treat volatility as a temporal cluster [2], TAGN conceptualizes crisis as a rapid reconfiguration of risk transmission channels, as evidenced by the sharp drop in Jaccard similarity (0.34) during the SVB crisis [39]. The observed 1–2 week lead time over the VIX index [40] validates that latent stress in supply chains (A_{supply}) [41] and institutional holdings (A_{hold}) [42] and manifests far earlier than broad market price volatility [40]. The distribution of realized volatility [43] and systemic risk measures in financial networks [44] further support these findings.

For regulatory bodies and RegTech applications, TAGN offers a distinct advantage in explainability [45]. By analyzing the attention weights (α_{ij}) generated by the GAT layer [46], regulators can identify “Hidden Hubs”—institutions that appear statistically insignificant during normal periods but become central nodes of contagion during stress events [46]. This capability facilitates a transition from reactive bailouts to precision intervention, allowing supervisors to monitor and potentially sever specific toxic transmission links identified through the dynamic graph evolution [47].

6.2. Limitations and Future Challenges

Despite its efficacy, the current iteration of the TAGN framework faces three primary practical challenges:

1. **Computational Complexity:** The dynamic attention mechanism scales linearly with the number of edges, $O(|E|)$ [48]. In a dense, full-market graph, the GPU memory consumption becomes a significant bottleneck [49]. This limitation, also discussed in the context of WinGNN’s gradient aggregation [4], suggests that future iterations must implement sparse matrix operations or graph sampling to maintain scalability for larger asset universes [50].
2. **Data Latency and Reporting Lags:** The institutional holding graph (A_{hold}) depends on Form 13F filings [51]. Because these are reported quarterly with a 45-day lag [51], it creates a temporal “blind spot”, where the model may fail to capture high-frequency position adjustments by hedge funds during rapidly evolving crises [52].
3. **Extreme Class Imbalance and Rare Events:** Financial crashes are, by definition, “Black Swan” events [53]—rare, high-impact occurrences that are difficult to predict using traditional statistical methods. While the use of Focal Loss with $\gamma = 2.0$ [54] partially mitigates the scarcity of positive samples ($y = 1$) [26], the model remains susceptible to overfitting stochastic noise in prolonged bull markets. Similarly to considerations in multimodal frameworks like DASNet [6], human expert validation remains essential for final high-stakes decision-making [55].

7. Conclusions and Future Work

This study presents the **Temporal Attentive Graph Networks (TAGN)**, a novel framework designed to bridge the gap between static network analysis and dynamic financial risk modeling. By *synergistically* coupling the spatial heterogeneity captured by GAT with the temporal memory of GRU, and integrating multi-scale heterogeneous graphs, TAGN effectively addresses the challenge of adapting to rapid topological evolution in financial markets.

Empirical evaluations conducted on the S&P 500 dataset (2018–2024) demonstrate that TAGN significantly outperforms state-of-the-art baselines, including PatchTST and WinGNN, in predictive accuracy. Notably, the model provides early-warning signals that lead the VIX index by approximately one to two weeks. This research advocates for a paradigm shift in systemic risk management: transitioning from traditional history-based statistical estimation toward a more proactive dynamic network-aware sensing approach.

To further enhance the robustness and applicability of the TAGN framework, future research will focus on the following three directions:

1. **LLM-Driven Causal Knowledge Graphs:** We plan to leverage Large Language Models (LLMs) to extract explicit causal chains from unstructured financial news (e.g., Chip Shortage to Auto Production Cut) [7]. By replacing statistical correlation graphs with Causal Knowledge Graphs, we aim to filter out spurious connections and enhance the model’s interpretability.

2. **Online Graph Learning for Non-Stationary Markets:** To address the challenge of *concept drift* in volatile financial environments, we intend to develop online learning algorithms [56]. This will enable TAGN to update its parameters in real-time as data streams arrive, ensuring the model remains adaptive without requiring frequent and costly offline retraining.
3. **Multiplex Network Modeling for Cross-Asset Spillovers:** We aim to expand the model's scope by constructing Multiplex Networks [57]. By representing the inter-dependencies between stock, bond, and foreign exchange markets as distinct yet interacting layers, the framework can better capture complex cross-asset risk spillover effects and systemic contagion.

Author Contributions: Conceptualization, W.Z.; methodology, W.Z.; software, W.Z., B.Z. and X.Z.; validation, W.Z.; formal analysis, W.Z.; investigation, W.Z.; resources, W.Z.; data curation, W.Z.; writing—original draft preparation, W.Z.; writing—review and editing, W.Z.; visualization, W.Z.; supervision, W.Z., Y.S., H.Z., and X.C.; project administration, W.Z., H.Z. and X.C.; funding acquisition, Y.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Sichuan Provincial Regional Innovation Cooperation Project (No. 2024YFHZ0085): “Joint Technical Research and Large-Scale Pilot Demonstration of Sichuan-Chongqing Data Elements Based on Unified Identifiers” and the Chengdu Science and Technology Program Project (No. 2023-YF11-00092-HZ): “Research, Development, and Application of Integrated Dynamic Mutual Recognition Technology for Sichuan-Chongqing Data Element Identifiers”.

Data Availability Statement: The data presented in this study are available in Zenodo at <https://zenodo.org/records/17785157> (DOI: 10.5281/zenodo.17785156) and in the public code repository on Gitee at https://gitee.com/Maxwell_Peng_pengjin/temporal-attentive-graph-networks-for-financial-surveillance.git. These data were derived from the following resources available in the public domain: SEC EDGAR database (<https://www.sec.gov/edgar>) and Yahoo Finance API (<https://finance.yahoo.com/>).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Elliott, M.; Golub, B.; Jackson, M. Financial networks and contagion. *Am. Econ. Rev.* **2014**, *104*, 3115–3153.
2. Bollerslev, T. Generalized autoregressive conditional heteroskedasticity. *J. Econom.* **1986**, *31*, 307–327.
3. Petrosino, L.; Bacco, L.; Salvati, G.; Merone, M. A GARCH-temporal fusion transformer model for the volatility prediction of exchange traded funds. *Neural Comput. Appl.* **2025**, *37*, 21435–21458.
4. Zhu, Y.; Cong, F.; Zhang, D.; Gong, W.; Lin, Q.; Feng, W.; Dong, Y.; Tang, J. WinGNN: Dynamic Graph Neural Networks with Random Gradient Aggregation Window. In Proceedings of the Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '23), Long Beach, CA, USA, 2023; pp. 3650–3662.
5. Guang, M.; Li, Z.; Yan, C.; Xu, Y.; Wang, J.; Cheng, D.; Jiang, C. Multi-Temporal Partitioned Graph Attention Networks for Financial Fraud Detection. *IEEE Trans. Inf. Forensics Secur.* **2025**, *20*, 142–157.
6. Nguyen, N.H.; Nguyen, T.T.; Ngo, Q.T. DASF-Net: A Multimodal Framework for Stock Price Forecasting with Diffusion-Based Graph Learning and Optimized Sentiment Fusion. *J. Risk Financ. Manag.* **2025**, *18*, 89.
7. Ning, K.; Pan, Z.; Liu, Y.; Jiang, Y.; Zhang, J.Y.; Rasul, K.; Schneider, A.; Ma, L.; Nevmyvaka, Y.; Song, D. TS-RAG: Retrieval-Augmented Generation based Time Series Foundation Models are Stronger Zero-Shot Forecaster. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), San Diego, CA, USA, 30 November–7 December 2025.
8. Mantegna, R. Hierarchical structure in financial markets. *Eur. Phys. J. B-Condens. Matter Complex Syst.* **1999**, *11*, 193–197.
9. Hsu, Y.L.; Tsai, Y.C.; Li, C.T. FinGAT: Financial Graph Attention Networks for Recommending Top-K Profitable Stocks. *IEEE Trans. Knowl. Data Eng.* **2025**, *37*, 1245–1259.
10. Duan, M.; Zheng, T.; Gao, Y.; Wang, G.; Feng, Z.; Wang, X. DGA-GNN: Dynamic Grouping Aggregation GNN for Fraud Detection. *AAAI Conf. Artif. Intell.* **2024**, *38*, 11820–11828.

11. Mehrabian, A.; Hoseinzade, E.; Mazloun, M.; Chen, X. Mamba Meets Financial Markets: A Graph-Mamba Approach for Stock Price Prediction. In Proceedings of the ICASSP 2025—2025 IEEE International Conference on Acoustics, Speech and Signal Processing, Hyderabad, India, 6–11 April 2025.
12. Lin, S.; Lin, W.; Wu, W.; Zhao, F.; Mo, R.; Zhang, H. SegRNN: Segment Recurrent Neural Network for Long-Term Time Series Forecasting. In Proceedings of the Proceedings of the International Conference on Learning Representations (ICLR), Vienna, Austria, 7–11 May 2024.
13. Wei, J.; Yang, S.; Cui, Z. Integrated GARCH-GRU in Financial Volatility Forecasting. *arXiv* **2025**, arXiv:2504.09380.
14. Romero Martinez, M.; Carmona Ibanez, P.; Martinez Vargas, J. Predicting Business Failure with the XGBoost Algorithm: The Role of Environmental Risk. *Sustainability* **2025**, *17*, 4948.
15. Kim, H.; Sohn, S. Support vector machines for default prediction of SMEs based on technology credit. *Eur. J. Oper. Res.* **2010**, *201*, 838–846.
16. Silva, E.; Castilho, D.; Pereira, A.; Brandao, H. A neural network based approach to support the Market Making strategies in High-Frequency Trading. In Proceedings of the 2014 International Joint Conference on Neural Networks (IJCNN), Beijing, China, 6–11 July 2014; pp. 845–852.
17. Qian, H.; Zhou, H.; Zhao, Q.; Chen, H.; Yao, H.; Wang, J.; Liu, Z.; Yu, F.; Zhang, Z.; Zhou, J. MDGNN: Multi-relational Dynamic Graph Neural Network for Comprehensive and Dynamic Stock Investment Prediction. *AAAI Conf. Artif. Intell.* **2024**, *38*, 8901–8909.
18. Huang, X.; Wang, J.; Lan, Y.; Jiang, C.; Yuan, X. MD-GCN: A multi-scale temporal dual graph convolution network for traffic flow prediction. *Sensors* **2023**, *23*, 841.
19. Zong, X.; Chen, Z.; Yu, F.; Wei, S. Local-global spatial-temporal graph convolutional network for traffic flow forecasting. *Electronics* **2024**, *13*, 636.
20. Chen, C.Y.; Huang, J.J. Temporal-guided knowledge graph-enhanced graph convolutional network for personalized movie recommendation systems. *Future Internet* **2023**, *15*, 323.
21. Ge, Y.; Wang, J.; Zhang, B.; Peng, F.; Ma, J.; Yang, C.; Zhao, Y.; Liu, M. Spatial-temporal-correlation-constrained dynamic graph convolutional network for traffic flow forecasting. *Mathematics* **2024**, *12*, 3159.
22. Tanvir, R.; Islam, M.; Sobhan, M.; Luo, D.; Mondal, A. MOGAT: A multi-omics integration framework using graph attention networks for cancer subtype prediction. *Int. J. Mol. Sci.* **2024**, *25*, 2788.
23. Wang, X.; Li, H.; Zhang, Z.; Chen, H.; Xiao, T.; Li, K.; Zhu, W. Uncertainty-aware Disentangled Dynamic Graph Attention Network for Out-of-Distribution Generalization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2025**, *47*, 1–18.
24. Xu, F.; Wang, N.; Wu, H.; Wen, X.; Zhao, X.; Wan, H. Revisiting Graph-based Fraud Detection in Sight of Heterophily and Spectrum. In Proceedings of the AAAI Conference on Artificial Intelligence, Vancouver, BC, Canada, 20–27 February 2024.
25. Chen, Z.; Zheng, L.; Lu, C.; Yuan, J.; Zhu, D. ChatGPT informed graph neural network for stock movement prediction. *arXiv* **2024**, arXiv:2401.05678.
26. He, H.; Garcia, E. Learning from imbalanced data. *IEEE Trans. Knowl. Data Eng.* **2009**, *21*, 1263–1284.
27. Ai, G.; Qiao, H.; Yan, H.; Pang, G. Semi-supervised Graph Anomaly Detection via Robust Homophily Learning. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), San Diego, CA, USA, 2–7 December 2025.
28. Lin, X.; Zhao, P.; Du, L. Edge-centric temporal GNNs for scalable fraud detection. *Adv. Neural Inf. Process. Syst.* **2023**, *36*, 2145–2157.
29. Han, Y.; Wang, L.; Cheng, Z.; Wang, B.; Yang, G.; Cheng, D.; Lin, X. Mitigating the Tail Effect in Fraud Detection by Community Enhanced Multi-Relation Graph Neural Networks. *IEEE Trans. Knowl. Data Eng.* **2025**, *37*, 502–515.
30. Choi, J.; Kim, H.; Whang, J.J. Unveiling the Threat of Fraud Gangs to Graph Neural Networks. In Proceedings of the AAAI Conference on Artificial Intelligence, Philadelphia, PA, USA, 25 February–4 March 2025.
31. Xu, J.; Perron, P. Forecasting carbon price: A novel multi-factor spatial-temporal GNN framework integrating Graph WaveNet and self-attention mechanism. *Energy Econ.* **2024**, *131*, 107385.
32. Xiang, S.; Cheng, D.; Shang, C.; Zhang, Y.; Liang, Y. Temporal and Heterogeneous Graph Neural Network for Financial Time Series Prediction. *arXiv* **2023**, arXiv:2305.08740.
33. Tang, Y.; Cai, Z. iTransformer-FFC: A Frequency-Aware Transformer Framework for Multi-Scale Time Series Forecasting. *Electronics* **2025**, *14*, 345.
34. Li, T.; Liu, Z.; Shen, Y.; Wang, X.; Chen, H.; Huang, S. MASTER: Market-Guided Stock Transformer for Stock Price Forecasting. *AAAI Conf. Artif. Intell.* **2024**, *38*, 11230–11238.
35. Nie, Y.; Nguyen, N.; Sinthong, P.; Kalagnanam, J. A Time Series is Worth 64 Words: Long-term Forecasting with Transformers. In Proceedings of the International Conference on Learning Representations (ICLR), Kigali, Rwanda, 1–5 May 2023.
36. Seo, Y.; Defferrard, M.; Vandergheynst, P.; Bresson, X. Structured sequence modeling with graph convolutional recurrent networks. In Proceedings of the International Conference on Neural Information Processing (ICONIP), Siem Reap, Cambodia, 13–16 December 2018; pp. 362–373.

37. Li, D.; Zhang, L.; Li, L. Forecasting stock volatility with economic policy uncertainty: A smooth transition GARCH-MIDAS model. *Int. Rev. Financ. Anal.* **2023**, *88*, 102660.
38. Engle, R.F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **1982**, *50*, 987–1007.
39. Silva, W. Network topology analysis of the US banking sector during the Silicon Valley Bank crisis. *J. Bank. Financ.* **2023**, *155*, 106921.
40. Fernandes, M. A two-factor model for the VIX index. *J. Bank. Financ.* **2014**, *46*, 272–283.
41. Barrot, J.; Sauvagnat, J. Input specificity and the propagation of idiosyncratic shocks in production networks. *Q. J. Econ.* **2016**, *131*, 1543–1592.
42. Anton, M.; Polk, C. Connected stocks. *J. Financ.* **2014**, *69*, 1099–1127.
43. Andersen, T.G.; Bollerslev, T.; Diebold, F.X.; Ebens, H. The distribution of realized stock return volatility. *J. Financ. Econ.* **2001**, *61*, 43–76.
44. Billio, M.; Getmansky, M.; Lo A.W.; Pelizzon, L. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. *J. Financ. Econ.* **2012**, *104*, 535–559.
45. Molnar, C. *Interpretable Machine Learning: A Guide for Making Black Box Models Explainable*; Independently Published, 2022. Available online: <https://christophm.github.io/interpretable-ml-book/> (accessed on 15 December 2025).
46. Velickovic, P.; Cucurull, G.; Casanova, A.; Romero, A.; Lio, P.; Bengio, Y. Graph attention networks. In Proceedings of the International Conference on Learning Representations (ICLR), Vancouver, BC, Canada, 30 April–3 May 2018.
47. Kazemi, S.M.; Goel, R.; Jain, K.; Kobayez, I.; Sethi, A.; Forsyth, P.; oupart, P. Representation learning for dynamic graphs: A survey. *J. Mach. Learn. Res.* **2020**, *21*, 1–73.
48. Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Li, Y.; Sun, M. Graph neural networks: A review of methods and applications. *AI Open* **2020**, *1*, 57–81.
49. Yang, J.; Tang, D.; Song, X.; Wang, L.; Yin, Q.; Chen, R.; Yu, W.; Zhou, J. GNNLab: A Factored System for Sample-based GNN Training over GPUs. In Proceedings of the 17th European Conference on Computer Systems (EuroSys '22), Rennes, France, 5–8 April 2022; pp. 417–434.
50. Hamilton, W.; Ying, Z.; Leskovec, J. Inductive representation learning on large graphs. In Proceedings of the Advances in Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 4–9 December 2017; Volume 30, pp. 1024–1034.
51. SEC. *Form 13F: Information Required of Institutional Investment Managers*; U.S. Securities and Exchange Commission: Washington, DC, USA, 2023.
52. Di Mascio, R. The information content of 13F filings. *J. Financ. Mark.* **2022**, *59*, 100742.
53. Taleb, N. *The Black Swan: The Impact of the Highly Improbable*; Random House: New York, NY, USA, 2007.
54. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 22–29 October 2017; pp. 2980–2988.
55. Liu, H.; Wang, Y.; Zhang, B.; Yao, B.; Liu, T.; Niu, Z.; Xue, X.; Tao, D.; Han, J.; Wang, W.; et al. Trustworthy AI: A computational perspective. *ACM Trans. Intell. Syst. Technol.* **2023**, *14*, 1–59.
56. He, Y. Continual learning in graph neural networks for evolving fraud patterns. In Proceedings of the International Conference on Learning Representations (ICLR), Vienna, Austria, 7–11 May 2024.
57. Wang, S.; et al. Stock Volatility Prediction via Bi-level Graph Attention. *arXiv* **2024**, arXiv:2409.12345.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.