MHG-FAN: Friction-Aware Heterogeneous Graphs for Stock Forecasting

Jiameng Chen, Zhongliang Yang*, Linna Zhou*
School of Cyberspace Security, Beijing University of Posts and Telecommunications, Beijing, China
yangzl@bupt.edu.cn

***Abstract*—Financial markets are driven by multimodal information streams where asset movements are shaped by complex intra- and inter-firm dependencies. However, traditional forecasting models struggle to characterize the complex mechanisms of risk transmission, often overlooking latent dependencies and the time-lagged nature of shock propagation. To address this challenge, we propose MHG-FAN, a risk-oriented framework that integrates multimodal market knowledge to construct a dynamic heterogeneous graph of financial entities. Specifically, we construct an adaptive dependency network that infers time-varying inter-firm relations from multimodal market signals, allowing the relational structure to evolve with changing market conditions. Building on this structure, we introduce a friction-aware relational module to model heterogeneous and asynchronous diffusion of market shocks, explicitly decoupling structural connectivity from transmission lags. Extensive experiments demonstrate that our approach consistently outperforms state-of-the-art baselines, offering an interpretable and effective tool for investment decision-making.**

***Index Terms*—Dynamic Heterogeneous Graphs, Stock Movement Prediction, Friction-Aware Modeling, Asynchronous Spillover**

I. Introduction

Stock movement prediction (SMP) aims to forecast the future price trend of listed companies and is a central problem in financial research and practice. Accurate prediction models can provide valuable guidance for investment strategies, risk control, and market regulation, yet the task remains highly challenging due to the stochastic nature of financial markets and the complex interplay of heterogeneous signals.

Traditional methods for SMP are typically divided into technical analysis and fundamental analysis. Technical analysis focuses on learning patterns from time-series signals such as prices and trading volume, using statistical tools or neural sequence models to capture temporal dependencies [1] [2]. While such methods can effectively exploit sequential features, they rely solely on individual stock histories and neglect broader market influences. Fundamental analysis, on the other hand, incorporates external information such as financial news, social media, or corporate disclosures to capture sentiment and event-driven signals [3] [4]. Despite their ability to enrich feature representations, these approaches generally treat each company in isolation and lack the capacity to reflect inter-stock dependencies.

Due to the momentum spillover effect in finance, stock prices are influenced not only by their own signals but also by the performance of related companies [5]. To capture such interdependence, graph-based methods have been proposed, where companies are modeled as nodes and relations such as industry categories, supply chains, or news co-occurrence are represented as edges. Variants of Graph Convolutional Networks (GCNs) [6], temporal extensions [7], and spatio-temporal hypergraph networks [8] have been developed to model cross-firm influences. Attribute-driven graph attention networks [9] further adjust predefined links by learning attribute-sensitive weights, while heterogeneous graph models incorporate multiple types of company relations [10] [11]. Despite these advances, most approaches still rely on static, predefined relations and fail to capture the dynamic evolution of market structures. They also overlook temporal lags in spillover effects, and insufficiently fuse multimodal signals, resulting in incomplete stock representations.

To address these challenges, we propose **MHG-FAN (Market Heterogeneous Graph Friction-Aware Network)**, a novel framework for stock movement prediction that jointly models multi-modal temporal signals and dynamic relational structures. Our approach captures both intra-firm sequential dependencies and inter-firm spillover effects through a unified architecture. The main contributions of our work are as follows:

- We introduce a dynamic graph learning mechanism that fuses explicit structural priors with implicit multimodal signals. This allows the model to continuously infer the market topology, capturing time-varying inter-stock dependencies that are invisible in static graphs.
- To overcome the limitation of synchronous message passing, we design a novel friction-aware lag mechanism. By modeling the anisotropic diffusion of market information, this module captures the delayed spillover effects caused by varying market friction, effectively resolving the spatio-temporal misalignment in risk propagation.
- Extensive empirical studies on two large-scale datasets (CSI100E and CSI300E) demonstrate that MHG-FAN significantly outperforms state-of-the-art baselines. Specifically, our model exhibits superior robustness in volatile market conditions, validating the necessity of modeling asynchronous mechanics.

II. Related Work

A. Sequence-Based Stock Prediction

Early studies model stock movement prediction using technical indicators derived from historical prices and trading volumes, with recurrent neural networks such as LSTM [1] and GRU [2] widely used to capture temporal dependencies. Jin et al. [12] further improve LSTM-based models by incorporating investor sentiment, multi-scale sequence decomposition, and attention mechanisms. To move beyond structured market signals, subsequent studies integrate fundamental and unstructured information into time-series models. Li et al. [13] associate news articles with related entities for stock movement forecasting, Wang et al. [14] leverage time-sensitive and target-aware investment opinions from social media with investor reliability modeling, and Yang et al. [4] exploit numeric-aware representations of earnings conference call data. Although these methods enrich single-stock temporal representations, they still model stocks independently and cannot capture evolving market-level relational dynamics.

B. Graph-Based Stock Prediction

To overcome the limitation of isolated stock modeling, graph neural networks have been introduced to capture inter-stock dependencies, with early studies applying GCN- or GAT-based architectures over predefined firm relations [15] [16]. Building on this idea, Sawhney et al. [17] integrate multimodal signals from prices, social media, and inter-stock correlations within a hierarchical temporal graph framework, while Cheng et al. [9] propose attribute-driven attention to adjust relation strengths for momentum spillover modeling. Luo et al. [18] explore cross-stock interactions from a causal and text-driven perspective using memory-aware graph mechanisms. To capture market heterogeneity, Zhao et al. [11] construct a bi-typed market knowledge graph involving companies and executives and learn momentum spillover via dual attention, and Peng et al. [19] enhance graph-based feature aggregation by introducing external attention and relation-aware aggregation strategies. Despite these advances, most graph-based approaches still rely on synchronous message passing over pre-defined or discretely updated relations, making it difficult to capture delayed and asymmetric spillover effects. We construct a dynamic market graph that integrates explicit structural priors with implicit multimodal signals and introduce a friction-aware lag mechanism to model asynchronous information diffusion.

III. Preliminaries

In the stock market, abundant signals are available, including historical trading data, news sentiment, and market relations, which together provide the basis for predicting stock movements. We study the stock movement prediction problem in a market with N listed companies. For each company $i \in \{1, \ldots, N\}$ and trading day t, we observe two types of sequential signals: the numerical indicators $\mathbf{p}_i^t \in \mathbb{R}^{d_m}$ derived from stock prices and volumes, and the textual sentiment features $\mathbf{q}_i^t \in \mathbb{R}^{d_n}$ extracted from financial news. We denote the fused daily representation as $\mathbf{x}_i^t = [\mathbf{p}_i^t \,\|\, \mathbf{q}_i^t]$, and over a lookback window of length T, the input sequence is $\mathbf{X}_i^{t-T+1:t} = \{\mathbf{x}_i^{t-T+1}, \ldots, \mathbf{x}_i^t\}$. In addition, inter-firm dependencies are encoded by a market graph $\mathcal{G}^t = (\mathcal{V}, \mathcal{E}^t)$, where $\mathcal{V}$ is the set of companies and $\mathcal{E}^t$ denotes both static relations (e.g., industry, supply chain, associations mediated by shared intermediaries) [11] and dynamic implicit relations inferred from market signals.

The prediction objective is to determine whether the stock price of company i will rise or fall on the next day $(t+1)$. Let c_i^t denote the closing price of stock i at day t. The binary movement label is defined as

$$y_i^{t+1} = \begin{cases} 1, & c_i^{t+1} \geq c_i^t, \\ 0, & c_i^{t+1} < c_i^t, \end{cases} \tag{1}$$

where $y_i^{t+1} = 1$ indicates an upward movement and $y_i^{t+1} = 0$ indicates a downward movement. Formally, the task is to learn a predictive function

$$f : \left(\mathbf{X}_i^{t-T+1:t}, \mathcal{G}^t\right) \rightarrow \hat{y}_i^{t+1}, \tag{2}$$

that maps multimodal historical signals and relational structures to the movement direction of each stock.

IV. Methodology

Our framework integrates multimodal sequential signals with dynamic relational structures to predict stock movements. As illustrated in Figure 1, MHG-FAN consists of three major components: (i) multimodal sequential encoding, (ii) asynchronous spillover via dynamic Heterogeneous Graph Modeling, and (iii) risk-oriented prediction layer. We describe each component in detail below.

A. Multimodal Sequential Encoding

Stock Representation Learning. For each company i at day t, we consider both numerical indicators $\mathbf{p}_i^t \in \mathbb{R}^{d_m}$ derived from stock trading data and sentiment signals $\mathbf{q}_i^t \in \mathbb{R}^{d_n}$ extracted from financial news. Specifically, $\mathbf{q}_i^t$ is obtained by aggregating all news articles related to company i that are published on day t. To mitigate noise and redundancy, we employ a low-dimensional and interpretable sentiment vector by computing lexicon-guided polarity and divergence scores at the article level and aggregating them into a daily representation. We first map the two modalities into a shared latent space through a tensor-based fusion function:

$$\mathbf{x}_i^t = \phi\big(\mathbf{W}_p \mathbf{p}_i^t + \mathbf{W}_q \mathbf{q}_i^t + \mathbf{T} \times (\mathbf{p}_i^t, \mathbf{q}_i^t)\big)\,, \tag{3}$$

where $\mathbf{W}_p$ and $\mathbf{W}_q$ are learnable projections, $\mathbf{T}$ is a tensor capturing cross-modal interactions, and $\phi(\cdot)$ is a nonlinear activation. This produces a unified daily representation $\mathbf{x}_i^t \in \mathbb{R}^{d_h}$.
Sequential modeling. Given a lookback window of length T, the multimodal sequence of company i is denoted as

$$\mathbf{X}_i^{t-T+1:t} = \{\mathbf{x}_i^{t-T+1}, \ldots, \mathbf{x}_i^t\}. \tag{4}$$

Since individual stocks exhibit distinct intrinsic volatilities and momentum patterns. Therefore, we employ a node-adaptive Gated Recurrent Unit (GRU) to model sequential dependencies. For each company i at time step τ, the hidden state is updated by node-specific gating functions:

$$\mathbf{h}_i^\tau = \text{GRUCell}\Big(\mathbf{x}_i^\tau, \mathbf{h}_i^{\tau-1}; \mathbf{W}^{(i)}, \mathbf{U}^{(i)}\Big)\,, \tag{5}$$

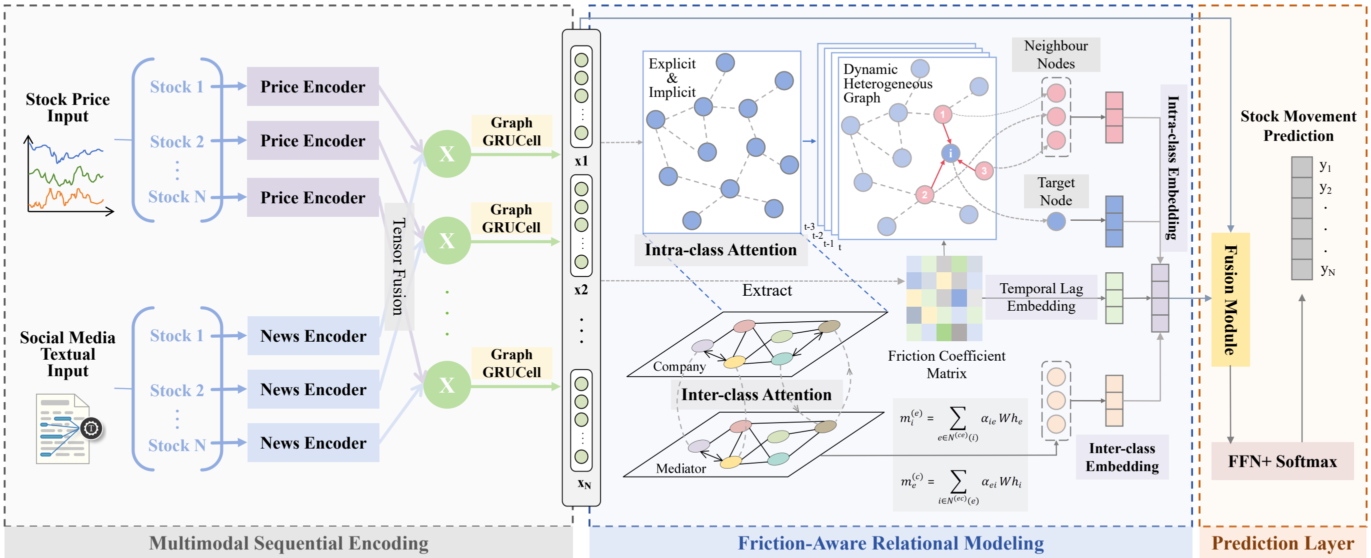

Fig. 1. **Overall architecture of the proposed MHG-FAN.** MHG-FAN fuses multimodal market signals to learn temporal embeddings and models explicit and implicit market relations through a heterogeneous graph network. A friction-aware lag module captures asynchronous spillover effects, and the fused representations are used for stock movement prediction.

where $\mathbf{x}_i^\tau$ is the fused input, $\mathbf{h}_i^{\tau-1}$ is the previous hidden state, and $\mathbf{W}^{(i)}, \mathbf{U}^{(i)}$ are gating weight matrices unique to node i, generated dynamically to capture firm-specific temporal behaviors.

Temporal attention. Financial signals naturally exhibit a "volatility decay" effect, where older information loses relevance. To mimic this internal market memory, we apply a time-decayed attention mechanism to aggregate the historical steps into the final sequential embedding:

$$\mathbf{s}_i^t = \sum_{\tau=t-T+1}^{t} \alpha_\tau \mathbf{h}_i^\tau, \quad \alpha_\tau = \frac{\exp(\beta_\tau \cdot \delta^{\,t-\tau})}{\sum_{k=t-T+1}^{t} \exp(\beta_k \cdot \delta^{\,t-k})}, \tag{6}$$

where β_τ is the raw attentional scoring scalar and $\delta \in (0, 1)$ serves as a learnable temporal decay coefficient, ensuring that recent shocks receive proportionally higher weights.

B. Friction-Aware Relational Modeling

Stock movements are shaped not only by intra-firm temporal patterns but also by complex inter-firm interactions, including indirect effects mediated by relational mediator nodes. To model these multi-scale dependencies, we construct a dynamic heterogeneous graph and design a friction-aware mechanism that captures (i) the evolving relational structure and (ii) the temporal lag of spillover effects.

Dynamic heterogeneous graph construction. At each time step t, we construct a heterogeneous market graph $\mathcal{G}^t = (\mathcal{V}, \mathcal{E}^t)$ with company nodes $\mathcal{V}_c$ and relational mediator nodes $\mathcal{V}_e$, where each node in $\mathcal{V}_e$ represents an association hub instantiated from available relational priors (e.g., shared affiliated entities or common-information exposure) and is connected to the set of companies it associates. To simulate realistic market topology, information flows along two complementary channels: (i) direct company–company edges, encompassing prior market relations (e.g., supply chain, industry) and dynamically inferred connections; and (ii) indirect inter-firm associations mediated by $\mathcal{V}_e$, which enable second-order propagation among firms linked by these association priors. This heterogeneous design captures both direct structural effects and higher-order relation-mediated interactions.

Multimodal-Driven Adaptive Topology. To capture inter-firm dependencies, we utilize both explicit structural priors (e.g., industry or supply chain relations) and adaptively inferred implicit edges. While the explicit graph encodes stable market knowledge, the implicit part complements it by actively discovering time-varying dependencies driven by real-time signals. At each time step t, the model dynamically infers the latent topology by computing the pairwise semantic affinity based on the multimodal sequential embeddings $\mathbf{s}^t$:

$$\tilde{\mathbf{A}}_{ij}^t = \frac{\exp\left(g([\mathbf{s}_i^t \,\|\, \mathbf{s}_j^t])\right)}{\sum_{k \neq i} \exp\left(g([\mathbf{s}_i^t \,\|\, \mathbf{s}_k^t])\right)}, \tag{7}$$

where $g(\cdot)$ is a learnable metric function projecting the concatenated multimodal features into an affinity score. The final adjacency matrix integrates both explicit and inferred relations as:

$$\mathbf{A}^t = \lambda \mathbf{A}^{\text{explicit}} + (1-\lambda)\tilde{\mathbf{A}}^t, \tag{8}$$

which allows the framework to jointly leverage predefined market knowledge and adaptively discovered dependencies.

Friction-Aware Temporal Lag Encoding. Standard GNNs aggregate spatial information synchronously. However, market spillover involves temporal delays governed by market friction. To explicitly model this asymmetric spillover, we introduce a friction-aware temporal encoding offset $\mathbf{o}_i^t$. By aggregating the historical neighbor states over a lookback window T, we apply an exponential decay to model the latency:

$$\mathbf{o}_i^t = \sum_{\tau=1}^{T} \exp(-\theta \cdot \tau) \cdot \left(\mathbf{A}_{i,:}^{t-\tau}\, \mathbf{s}^{t-\tau}\right), \tag{9}$$

where $\theta > 0$ is a learnable market friction coefficient. A larger θ implies high friction (rapid decay of past shocks), while a smaller θ preserves the long-term spillover footprint.

Asynchronous Spatio-Temporal Aggregation. Given the heterogeneous graph $\mathcal{G}^t$, each company's representation is updated by asynchronously aggregating direct and mediated signals.

For company i, immediate messages from direct market neighbors are computed via an attention-weighted transformation:

$$\mathbf{m}_i^{(c)} = \sigma\left(\sum_{j\in\mathcal{N}_i^{(c)}} \alpha_{ij}^{(c)} \mathbf{W}_c \mathbf{s}_j^t\right), \tag{10}$$

where $\alpha_{ij}^{(c)}$ are normalized attention scores and $\sigma(\cdot)$ is a nonlinear activation. Here, $\mathbf{s}_j^t$ is used as the node feature.

To capture systemic risk spillover through relation-mediated channels, messages are also propagated via relational mediator nodes, forming a 2-hop message passing process:

(1) **Aggregation (Company $\rightarrow$ Relational Mediator):** The latent state of relational mediator node k at time t is updated by aggregating the sequential momentum of its associated companies $\mathcal{N}_k^{(ec)}$ and peer influences from other relational mediator nodes $\mathcal{N}_k^{(ee)}$:

$$\mathbf{e}_k^t = \phi\left(\sum_{j\in\mathcal{N}_k^{(ec)}} \alpha_{kj}^{(ec)} \mathbf{W}_{ec} \mathbf{s}_j^t + \sum_{l\in\mathcal{N}_k^{(ee)}} \alpha_{kl}^{(ee)} \mathbf{W}_{ee} \mathbf{e}_l^{t-1}\right), \tag{11}$$

where $\mathbf{e}_k^t$ denotes the aggregated state of mediator node k at time t, and α are normalized relational attention weights.

(2) **Dispersal (Relational Mediator $\rightarrow$ Company):** The aggregated information is then transmitted back to company i from its associated relational mediator nodes $\mathcal{N}_i^{(ce)}$:

$$\mathbf{m}_i^{(e)} = \phi\left(\sum_{k\in\mathcal{N}_i^{(ce)}} \alpha_{ik}^{(ce)} \mathbf{W}_{ce} \mathbf{e}_k^t\right). \tag{12}$$

By explicitly modeling this Company–Mediator–Company path, $\mathbf{m}_i^{(e)}$ captures second-order momentum spillover beyond explicit market links.

Finally, to resolve the spatio-temporal misalignment, the representation is obtained by fusing $\mathbf{m}_i^{(c)}$ and $\mathbf{m}_i^{(e)}$, and correcting them with the temporal lag offset $\mathbf{o}_i^t$:

$$\mathbf{h}_i^{\text{rel}} = \text{LayerNorm}\left(\mathbf{m}_i^{(c)} + \mathbf{m}_i^{(e)}\right) + \mathbf{o}_i^t. \tag{13}$$

By unifying intra-firm dynamics and inter-firm asymmetric spillover, the relational embedding $\mathbf{h}_i^{\text{rel}}$ reflects time-varying dependencies among companies and relational mediator nodes.

C. Risk-Oriented Prediction and Optimization

State Fusion. We fuse the intra-firm sequential embedding $\mathbf{s}_i^t$ and the inter-firm asynchronous relational embedding $\mathbf{h}_i^{\text{rel}}$ via a learnable gating mechanism, selectively balancing internal alpha and external risk spillover:

$$\mathbf{g}_i^t = \sigma\big(\mathbf{W}_g[\mathbf{s}_i^t \| \mathbf{h}_i^{\text{rel}}]\big), \quad \mathbf{z}_i^t = \mathbf{g}_i^t \odot \mathbf{s}_i^t + (1-\mathbf{g}_i^t) \odot \mathbf{h}_i^{\text{rel}}. \tag{14}$$

Prediction Head and Optimization. The fused representation $\mathbf{z}_i^t$ is projected via a Multilayer Perceptron (MLP) to output the final directional probability:

$$\hat{y}_i^{t+1} = \text{Softmax}\big(\mathbf{W}_2, \phi(\mathbf{W}_1 \mathbf{z}_i^t + \mathbf{b}_1) + \mathbf{b}_2\big). \tag{15}$$

Aligned with our risk-oriented philosophy, the model is trained end-to-end to minimize the binary cross-entropy loss, fortified by an entropy regularization term $\mathcal{L}_{\text{attn}}$ to ensure stability in the dynamic graph distribution during extreme market conditions:

$$\mathcal{L} = -\sum_{i,t} y_i^{t+1} \log \hat{y}i^{t+1} + \eta \mathcal{L}\text{attn}. \tag{16}$$

V. Experiments

A. Experimental settings

Datasets. We conduct experiments on two open-source stock market datasets, CSI100E and CSI300E [11], both derived from the China Securities Index and covering daily trading data and financial news. CSI100E includes 73 companies, while CSI300E contains 185. Each company is described by numerical indicators and sentiment features extracted from major financial media. The datasets also provide multiple relation types among companies and association entities, forming a bi-typed market knowledge graph. The observation period spans November 21, 2017 to December 31, 2019, with data split into training (Nov 2017–Aug 2019), validation (Aug–Oct 2019), and test (Oct–Dec 2019) sets.

Baselines. We compare our model against a comprehensive set of baselines covering sequential models, graph-based models, and finance-specific graph extensions. For sequential modeling, we include classical recurrent neural networks such as LSTM [1] and GRU [2]. For graph representation learning, we evaluate standard architectures including GCN [15], GAT [16], RGCN [20], and HGT [21], as well as advanced variants such as Multi-GCGRU [22], MAN-SF [17], and STHAN-SR [8] that integrate temporal or multi-graph information. In addition, we consider domain-specific methods designed for stock movement prediction, including AD-GAT [9], DANSMP [11], EA-RTGCN [19] and MTIDHHGAN [23]. These baselines represent a broad spectrum of approaches, from generic deep learning models to specialized graph-based techniques for financial applications.

Metrics. We evaluate prediction performance using two widely adopted metrics: directional accuracy (DA) and the area under the ROC curve (AUC).

DA measures the proportion of samples whose predicted movement direction matches the actual trend:

$$DA = \frac{n}{N}, \tag{17}$$

where n is the number of correct predictions and N is the total number of predictions.
AUC evaluates the ranking ability of the model by measuring the probability that a randomly chosen positive sample receives a higher predicted score than a negative one:

$$\mathrm{AUC} = \frac{1}{N^+N^-} \sum_{i:y_i=1} \sum_{j:y_j=0} \mathbb{I}(\hat{p}_i > \hat{p}_j), \tag{18}$$

where N^+ and N^- denote the numbers of positive and negative samples, respectively, and $\hat{p}_i$ is the predicted probability.

VI. Results and Evaluations

A. Comparison with Baselines

We first compare MHG-FAN with a wide range of baseline models, including sequential networks, standard graph models, advanced temporal or multi-graph extensions, and finance-oriented approaches. The results are summarized in Table I.

TABLE I
Performance comparison of different models on CSI100E and CSI300E datasets.

Models	CSI100E		CSI300E	
	DACC	AUC	DACC	AUC
LSTM [1]	51.14	51.33	51.78	52.24
GRU [2]	51.66	51.46	51.11	52.30
GCN [15]	51.58	52.18	51.68	51.81
GAT [16]	52.17	52.24	51.40	52.24
RGCN [20]	52.33	52.69	51.79	52.59
HGT [21]	53.01	52.51	51.70	52.19
MAN-SF [17]	52.86	52.23	51.19	52.48
Multi-GCGRU [22]	53.65	54.28	52.77	54.54
STHAN-SR [8]	52.78	53.05	52.89	53.48
AD-GAT [9]	54.56	55.46	52.63	54.29
DANSMP [11]	57.75	60.78	55.79	59.36
EA-RTGCN [19]	60.66	64.62	58.28	61.31
MTIDHHGAN [23]	61.37	64.44	56.48	60.18
Ours	**62.74**	**65.52**	**60.01**	**63.19**

Our model consistently achieves the best performance on both CSI100E and CSI300E datasets [11]. In particular, it reaches **62.74%** DA and **65.52%** AUC on CSI100E, and **60.01%** DA and **63.19%** AUC on CSI300E, outperforming the strongest baselines EA-RTGCN [19] and MTIDHHGAN [23]. The improvements are especially pronounced in AUC, indicating that our method not only predicts the movement direction accurately but also produces more reliable probability rankings. These results demonstrate the effectiveness of integrating sequential dynamics with dual-dynamics relational modeling.

B. Ablation Study

To further investigate the contribution of each component, we conduct ablation experiments on CSI100E, as reported in Table II.

TABLE II
Ablation study results of the effect of different modules.

Model	DACC	AUC
w/o Adaptive Implicit Topology	57.04	60.66
w/o Explicit Structural Priors	60.14	62.22
w/o Friction-Aware Lag Mechanism	59.78	62.15
w/o Relational Mediator Nodes	59.01	61.09
w/o Multimodal Fusion Module	58.85	60.01
Ours	**62.74**	**65.52**

The ablation results show that removing the adaptive implicit topology notably degrades performance, confirming the need for adaptive relation inference. Excluding explicit structural priors also degrades results, highlighting the importance of structural priors in guiding dynamic updates. The friction-aware lag mechanism improves both DA and AUC by modeling temporal spillovers, while removing relational mediator nodes weakens indirect interaction modeling. Finally, eliminating the multimodal fusion module reduces performance, highlighting the importance of combining sequential and relational embeddings. Overall, each component contributes to the final performance, and their integration achieves the best results.

C. Input Modality Analysis

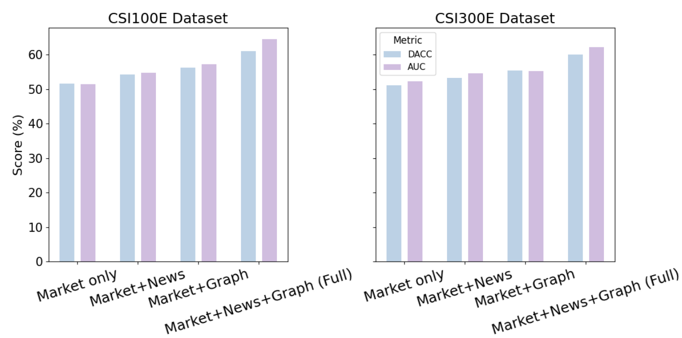

Fig. 2. Modality Contribution Results on CSI100E and CSI300E

To evaluate the contribution of different input modalities, we compare the performance of models with market data alone, market combined with news, market combined with graph information, and the full model integrating all three modalities. As shown in Figure 2, the results present that market features serve as the fundamental predictive signal, while the addition of news sentiment and graph structure each bring further improvements. Notably, the full integration of all modalities achieves the best overall performance, confirming their complementary roles in enhancing stock movement prediction. This demonstrates the advantage of our proposed multi-modal framework in effectively capturing diverse market signals for more accurate prediction.

D. Profitability and Robustness Analysis

We compare the cumulative returns of MHG-FAN with the market benchmark over a 50-day backtesting period in Figure 3, while Table III reports the corresponding risk–return metrics. As shown in Figure 3, MHG-FAN exhibits stronger downside resilience during market downturns, maintaining more stable performance and smaller drawdowns than the benchmark, which is consistent with the lower maximum drawdown in Table III. During market recovery and upward phases, MHG-FAN achieves steeper cumulative return growth, resulting in substantially higher annualized and total returns. Although the strategy exhibits higher volatility, its improved Sharpe ratio indicates superior risk-adjusted performance.

TABLE III
BACKTESTING METRICS COMPARISON BETWEEN OUR METHOD AND THE MARKET INDEX.

Metric	**Market Index**	**Ours**	**Improv.**
ARR (↑)	0.4323	0.8672	+100.6%
AV (↓)	0.1718	0.2327	-35.4%
ASR (↑)	2.0925	2.6874	+28.4%
MDD (↑)	-0.0625	-0.0573	+8.3%
Total Return (↑)	0.0693	0.1232	+77.8%

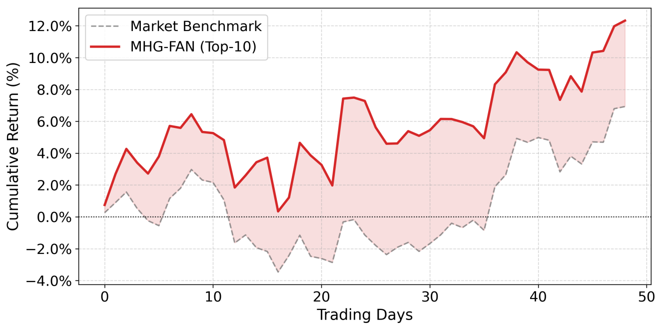

Fig. 3. Cumulative return comparison between MHG-FAN (Top-10 strategy) and the market index.

VII. Conclusion

In this paper, we propose MHG-FAN, a multimodal framework for stock movement prediction that jointly models temporal signals and complex relational structures. We construct an adaptive dependency network to capture latent inter-firm correlations and introduce a friction-aware mechanism to model asynchronous spillover effects of market shocks. Extensive experiments on CSI100E and CSI300E demonstrate that MHG-FAN consistently outperforms state-of-the-art baselines.

Acknowledgment

This work was supported in part by the National Key Research and Development Program of China under Grant 2023YFC3305402 and in part by the National Natural Science Foundation of China under Grant 62172053 and Grant 62302059.

References

[1] S. Hochreiter and J. Schmidhuber, “Long short-term memory,” *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[2] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” *arXiv preprint arXiv:1406.1078*, 2014.
[3] Y. Xu and S. B. Cohen, “Stock movement prediction from tweets and historical prices,” in *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 2018, pp. 1970–1979.
[4] L. Yang, J. Li, R. Dong, Y. Zhang, and B. Smyth, “Numhtml: Numeric-oriented hierarchical transformer model for multi-task financial forecasting,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 36, no. 10, 2022, pp. 11 604–11 612.
[5] U. Ali and D. Hirshleifer, “Shared analyst coverage: Unifying momentum spillover effects,” *Journal of Financial Economics*, vol. 136, no. 3, pp. 649–675, 2020.
[6] Y. Chen, Z. Wei, and X. Huang, “Incorporating corporation relationship via graph convolutional neural networks for stock price prediction,” in *Proceedings of the 27th ACM international conference on information and knowledge management*, 2018, pp. 1655–1658.
[7] F. Feng, X. He, X. Wang, C. Luo, Y. Liu, and T.-S. Chua, “Temporal relational ranking for stock prediction,” *ACM Transactions on Information Systems (TOIS)*, vol. 37, no. 2, pp. 1–30, 2019.
[8] R. Sawhney, S. Agarwal, A. Wadhwa, T. Derr, and R. R. Shah, “Stock selection via spatiotemporal hypergraph attention network: A learning to rank approach,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 1, 2021, pp. 497–504.
[9] R. Cheng and Q. Li, “Modeling the momentum spillover effect for stock prediction via attribute-driven graph attention networks,” in *Proceedings of the AAAI conference on artificial intelligence*, vol. 35, no. 1, 2021, pp. 55–62.
[10] S. Xiang, D. Cheng, C. Shang, Y. Zhang, and Y. Liang, “Temporal and heterogeneous graph neural network for financial time series prediction,” in *Proceedings of the 31st ACM international conference on information & knowledge management*, 2022, pp. 3584–3593.
[11] Y. Zhao, H. Du, Y. Liu, S. Wei, X. Chen, F. Zhuang, Q. Li, and G. Kou, “Stock movement prediction based on bi-typed hybrid-relational market knowledge graph via dual attention networks,” *IEEE transactions on knowledge and data engineering*, vol. 35, no. 8, pp. 8559–8571, 2022.
[12] Z. Jin, Y. Yang, and Y. Liu, “Stock closing price prediction based on sentiment analysis and lstm,” *Neural Computing and Applications*, vol. 32, no. 13, pp. 9713–9729, 2020.
[13] W. Li, R. Bao, K. Harimoto, D. Chen, J. Xu, and Q. Su, “Modeling the stock relation with graph network for overnight stock movement prediction,” in *Proceedings of the twenty-ninth international conference on international joint conferences on artificial intelligence*, 2021, pp. 4541–4547.
[14] H. Wang, T. Wang, and Y. Li, “Incorporating expert-based investment opinion signals in stock prediction: A deep learning framework,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 01, 2020, pp. 971–978.
[15] T. Kipf, “Semi-supervised classification with graph convolutional networks,” *arXiv preprint arXiv:1609.02907*, 2016.
[16] V. Petar, C. Guillem, C. Arantxa, R. Adriana, L. Pietro, and B. Yoshua, “Graph attention networks,” in *International conference on learning representations*, vol. 8, 2018.
[17] R. Sawhney, S. Agarwal, A. Wadhwa, and R. Shah, “Deep attentive learning for stock movement prediction from social media text and company correlations,” in *Proceedings of the 2020 conference on empirical methods in natural language processing (EMNLP)*, 2020, pp. 8415–8426.
[18] D. Luo, S. Li, W. Liao, and R. Yan, “Magicnet: Memory-aware graph interactive causal network for multivariate stock price movement prediction,” *IEEE Transactions on Knowledge and Data Engineering*, 2025.
[19] H. Peng, K. Dong, and J. Yang, “Stock price movement prediction based on relation type guided graph convolutional network,” *Engineering Applications of Artificial Intelligence*, vol. 126, p. 106948, 2023.
[20] M. Schlichtkrull, T. N. Kipf, P. Bloem, R. Van Den Berg, I. Titov, and M. Welling, “Modeling relational data with graph convolutional networks,” in *European semantic web conference*. Springer, 2018, pp. 593–607.
[21] Z. Hu, Y. Dong, K. Wang, and Y. Sun, “Heterogeneous graph transformer,” in *Proceedings of the web conference 2020*, 2020, pp. 2704–2710.
[22] J. Ye, J. Zhao, K. Ye, and C. Xu, “Multi-graph convolutional network for relationship-driven stock movement prediction,” in *2020 25th International Conference on Pattern Recognition (ICPR)*. IEEE, 2021, pp. 6702–6709.
[23] Z. Wu, P. Chen, J. Yu, and X. Ye, “Mtidhhgan: A multimodal temporal information-driven hierarchical heterogeneous graph attention network for stock movement prediction,” in *Proceedings of the 2025 4th International Conference on Cyber Security, Artificial Intelligence and the Digital Economy*, 2025, pp. 322–330.