DySTAGE: Dynamic Graph Representation Learning for Asset Pricing via Spatio-Temporal Attention and Graph Encodings

Speaker: Jingyi Gu (New Jersey Institute of Technology)

Authors: Jingyi Gu, Junyi Ye, Ajim Uddin, Grace Wang

2024/11/16

Asset pricing in financial networks

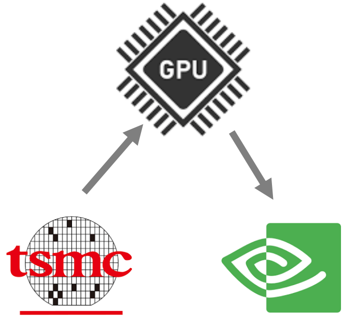
Task: estimate future values/returns of financial assets by modeling their interactions

- Assets are interdependent through various dimensions

Asset pricing in financial networks

Task: estimate future values/returns of financial assets by modeling their interactions

- Assets are interdependent through various dimensions

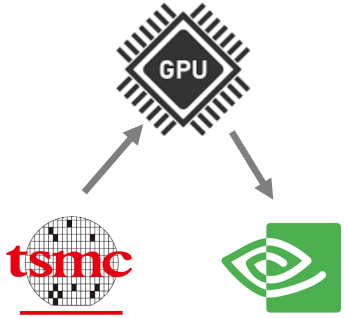


Supply chain
TSMC vs Nvidia

Asset pricing in financial networks

Task: estimate future values/returns of financial assets by modeling their interactions

- Assets are interdependent through various dimensions



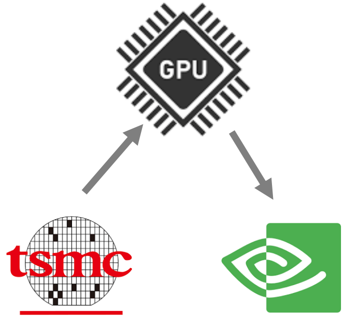
Supply chain
TSMC vs Nvidia

Industry sector
Coca-cola vs Pepsi

Asset pricing in financial networks

Task: estimate future values/returns of financial assets by modeling their interactions

- Assets are interdependent through various dimensions



Supply chain
TSMC vs Nvidia

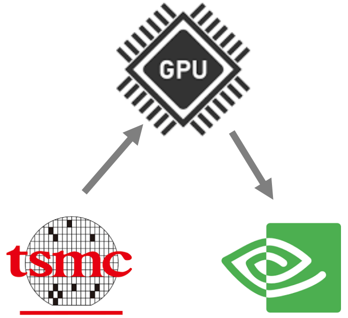
Industry sector
Coca-cola vs Pepsi

Macroeconomic conditions
Real estate vs banking

Asset pricing in financial networks

Task: estimate future values/returns of financial assets by modeling their interactions

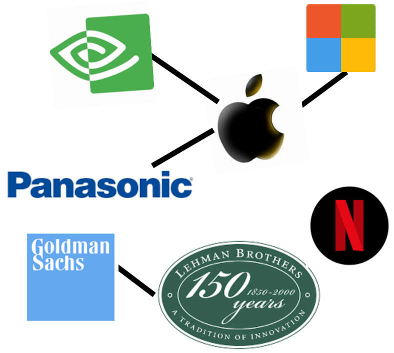
- Assets are interdependent through various dimensions



Supply chain
TSMC vs Nvidia

Industry sector
Coca-cola vs Pepsi

Macroeconomic conditions
Real estate vs banking

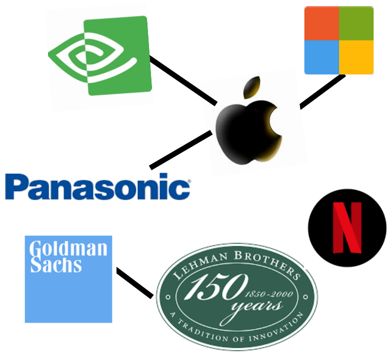


Graph Networks as a preferred tool

Dynamics in financial networks is largely ignored

Existing asset pricing models:

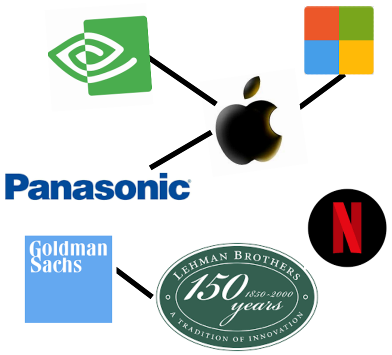
- fixed group of assets
- static relationship



Dynamics in financial networks is largely ignored

Existing asset pricing models:

- fixed group of assets
- static relationship

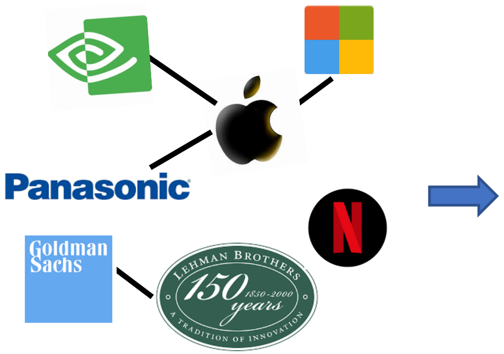


Financial networks in the real world evolve continuously

- technical innovations
- corporate events

- Changing composition of assets
- Evolving interrelationships

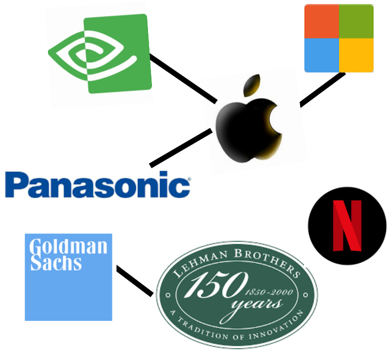
Market @ 2008-09



Dynamics in financial networks is largely ignored

Existing asset pricing models:

- fixed group of assets
- static relationship



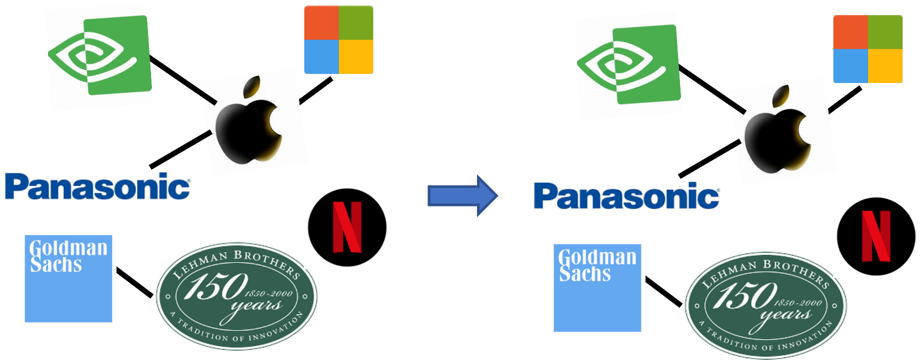
Financial networks in the real world evolve continuously

- technical innovations
- corporate events

- Changing composition of assets
- Evolving interrelationships

Market @ 2008-09

Market @ 2010-06



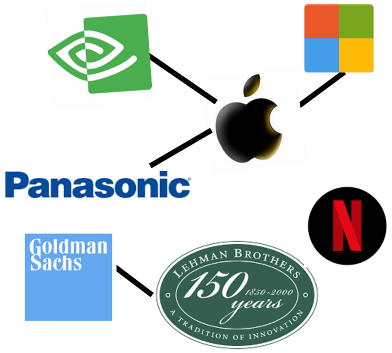
Lehman Brothers is broken

Tesla enter the market

Dynamics in financial networks is largely ignored

Existing asset pricing models:

- fixed group of assets
- static relationship



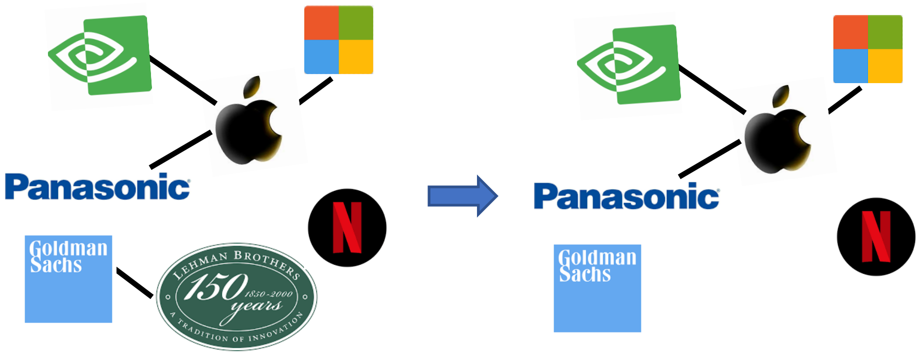
Financial networks in the real world evolve continuously

- technical innovations
- corporate events

- Changing composition of assets
- Evolving interrelationships

Market @ 2008-09

Market @ 2010-06

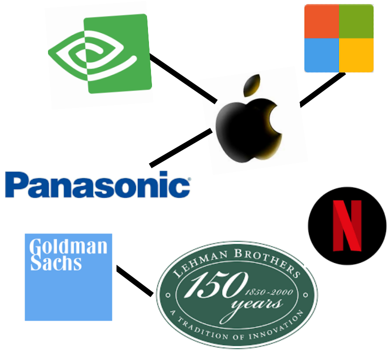


Lehman Brothers is broken
Tesla enter the market

Dynamics in financial networks is largely ignored

Existing asset pricing models:

- fixed group of assets
- static relationship



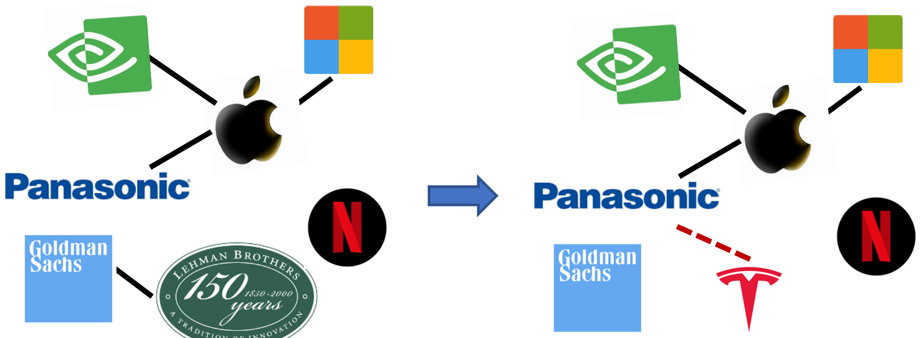
Financial networks in the real world evolve continuously

- technical innovations
- corporate events

- Changing composition of assets
- Evolving interrelationships

Market @ 2008-09

Market @ 2010-06

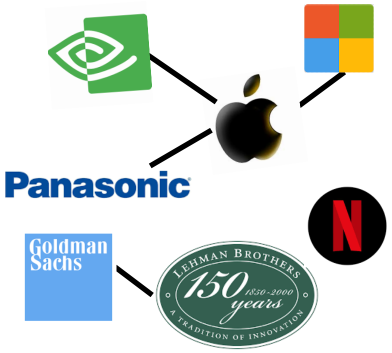


Lehman Brothers is broken
Tesla enter the market

Dynamics in financial networks is largely ignored

Existing asset pricing models:

- fixed group of assets
- static relationship

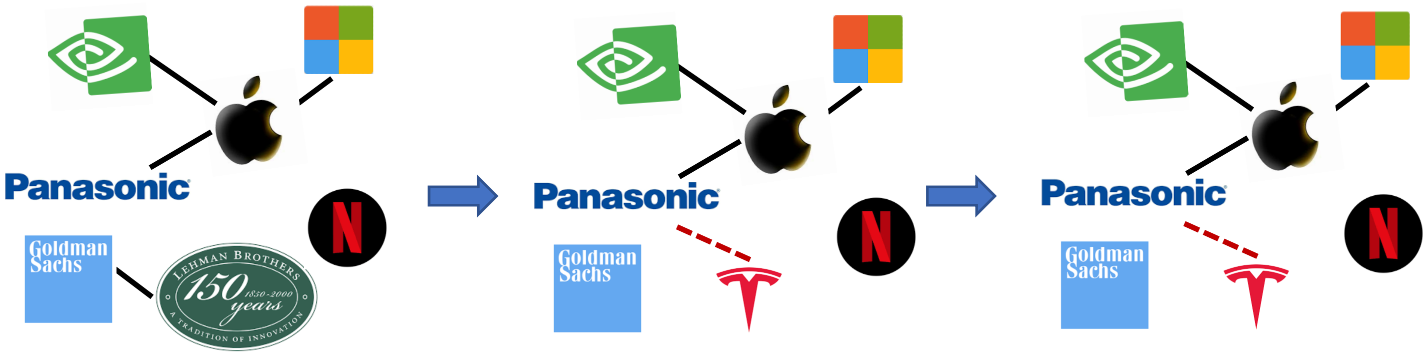


Financial networks in the real world evolve continuously

- technical innovations
- corporate events

- Changing composition of assets
- Evolving interrelationships

Market @ 2008-09 Market @ 2010-06 Market @ 2020-06



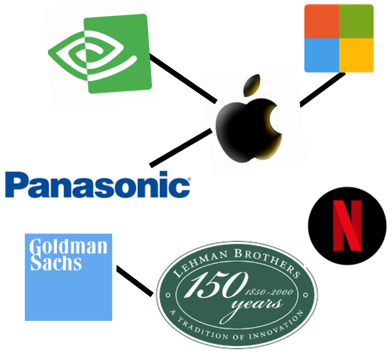
Lehman Brothers is broken
Tesla enter the market

Apple enters the streaming arena with the launch of Apple TV+

Dynamics in financial networks is largely ignored

Existing asset pricing models:

- fixed group of assets
- static relationship



Financial networks in the real world evolve continuously

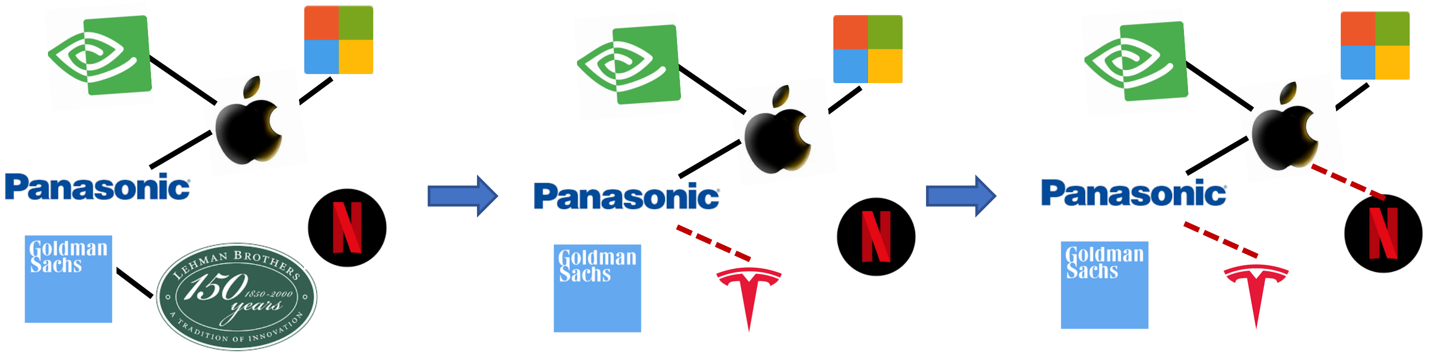
- technical innovations
- corporate events

- Changing composition of assets
- Evolving interrelationships

Market @ 2008-09

Market @ 2010-06

Market @ 2020-06



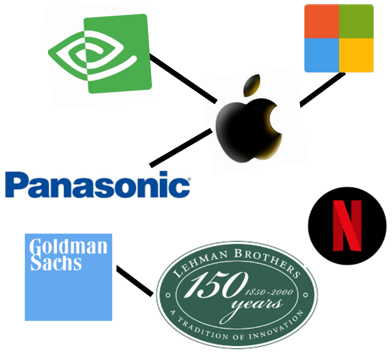
Lehman Brothers is broken
Tesla enter the market

Apple enters the streaming arena with the launch of Apple TV+

Dynamics in financial networks is largely ignored

Existing asset pricing models:

- fixed group of assets
- static relationship

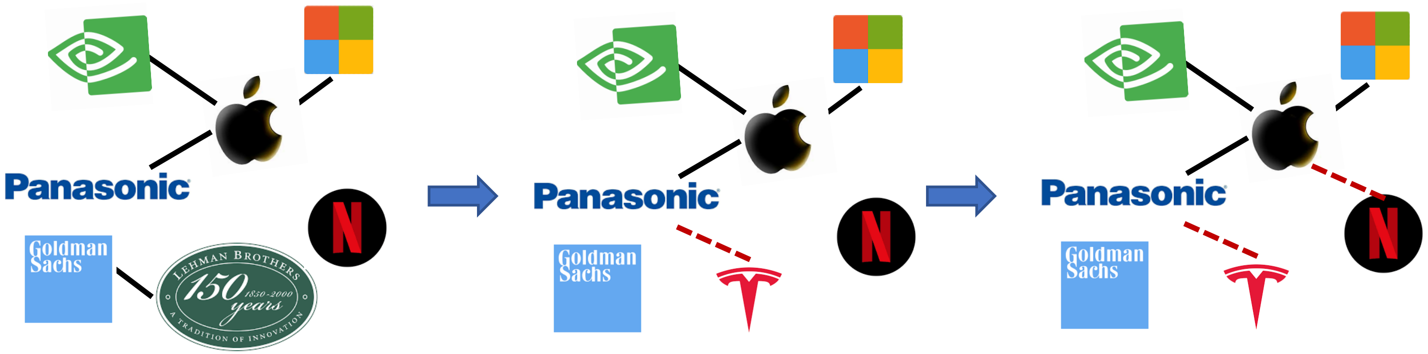


Financial networks in the real world evolve continuously

- technical innovations
- corporate events

- Changing composition of assets
- Evolving interrelationships

Market @ 2008-09 | Market @ 2010-06 | Market @ 2020-06

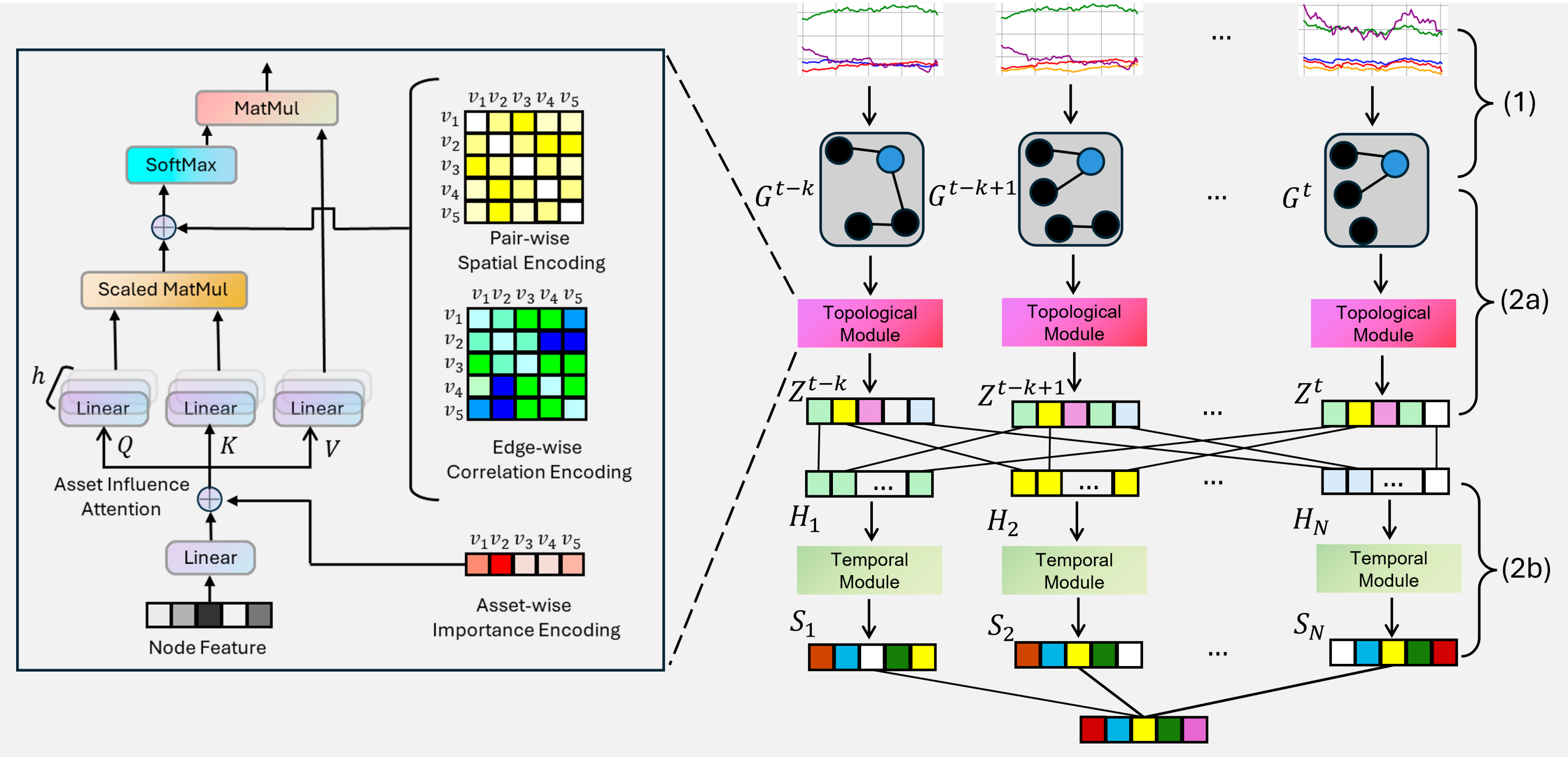


Lehman Brothers is broken
Tesla enter the market

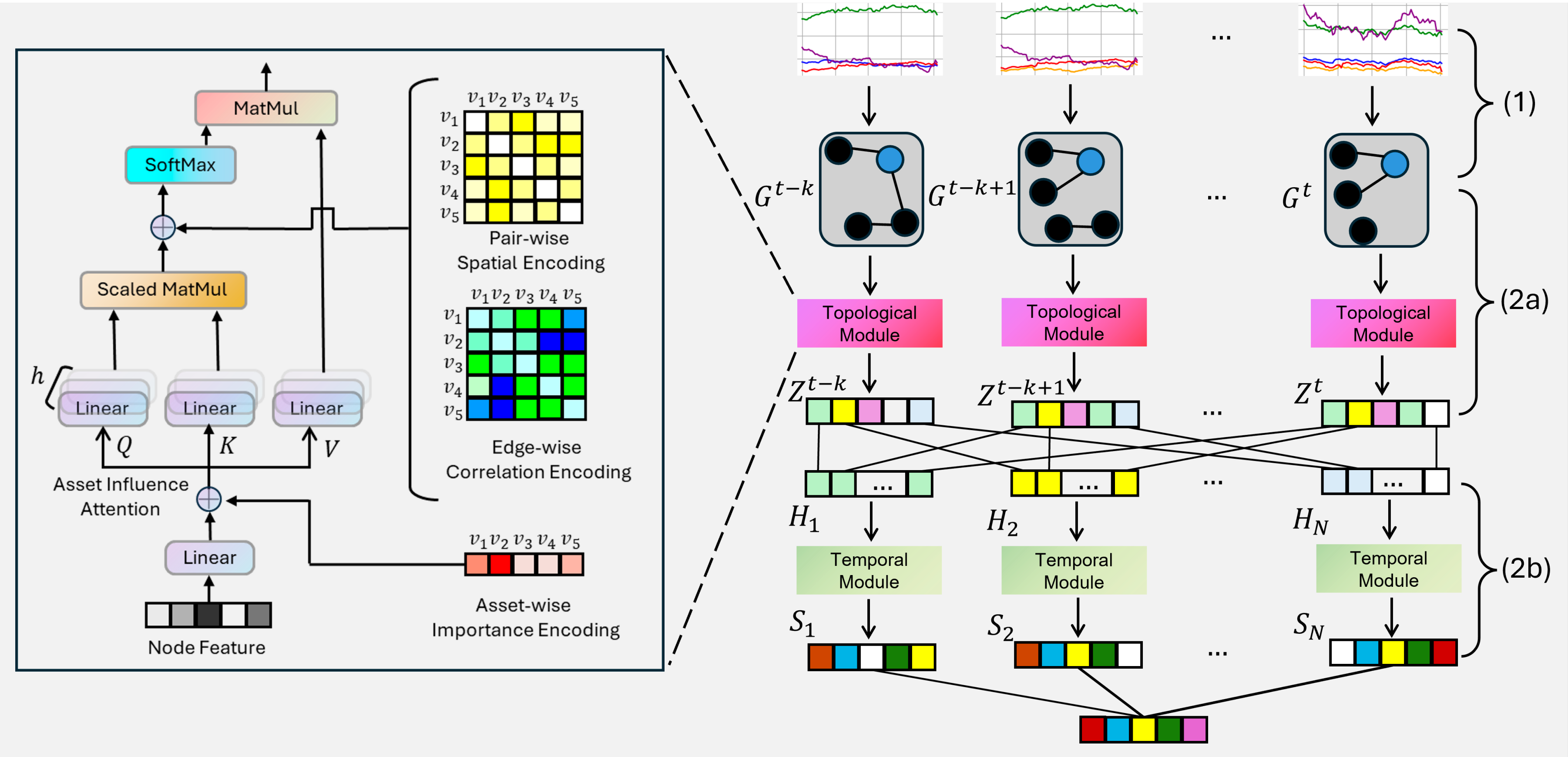
Apple enters the streaming arena with the launch of Apple TV+

A framework that represents time-varying dynamics of financial markets is necessary

Dynamic graph representation learning framework via Spatio-Temporal Attention and Graph Encodings (DySTAGE)

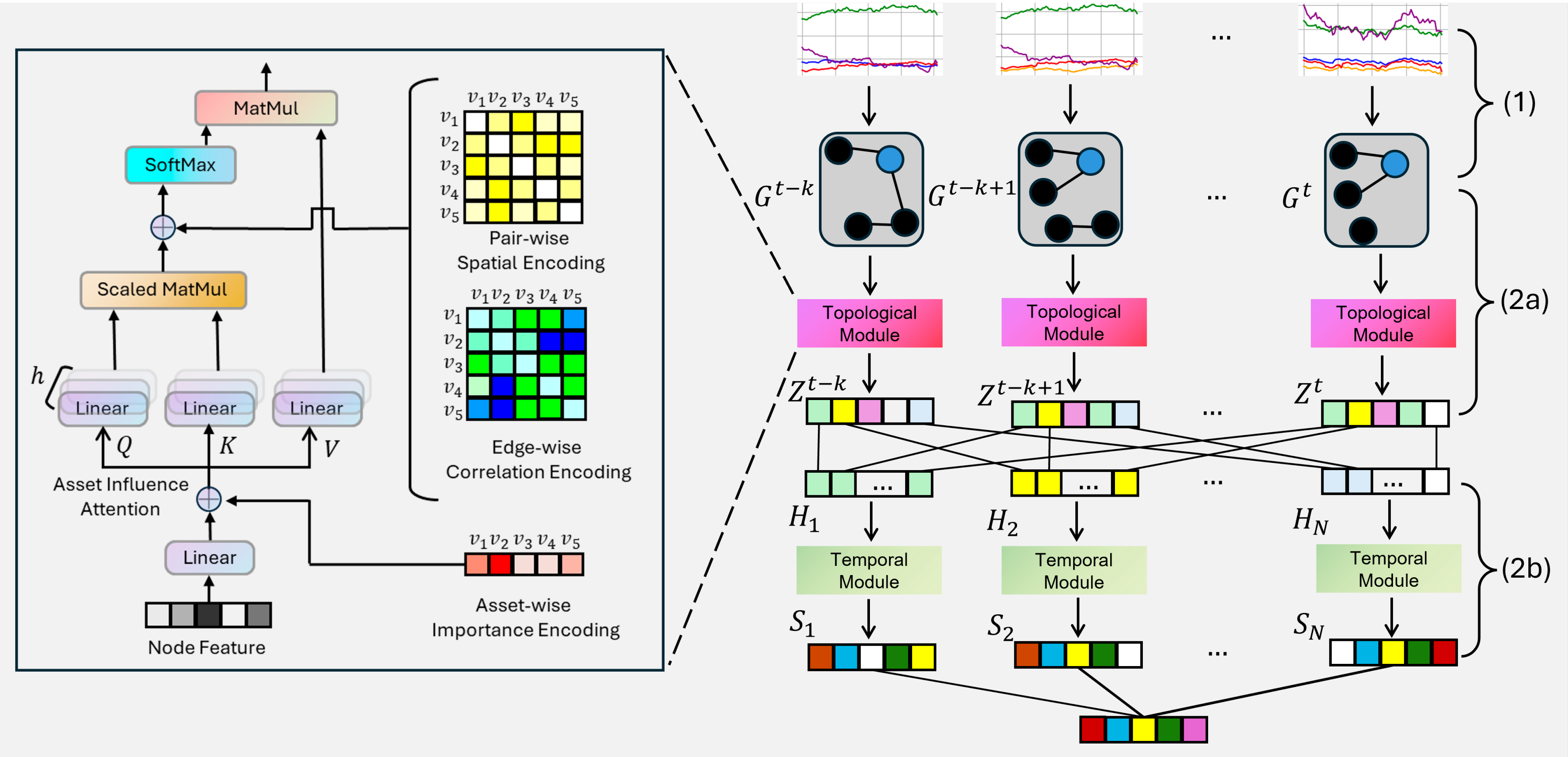


Dynamic graph representation learning framework via Spatio-Temporal Attention and Graph Encodings (DySTAGE)



A universal formulation for dynamic graph construction in asset pricing

Dynamic graph representation learning framework via Spatio-Temporal Attention and Graph Encodings (DySTAGE)

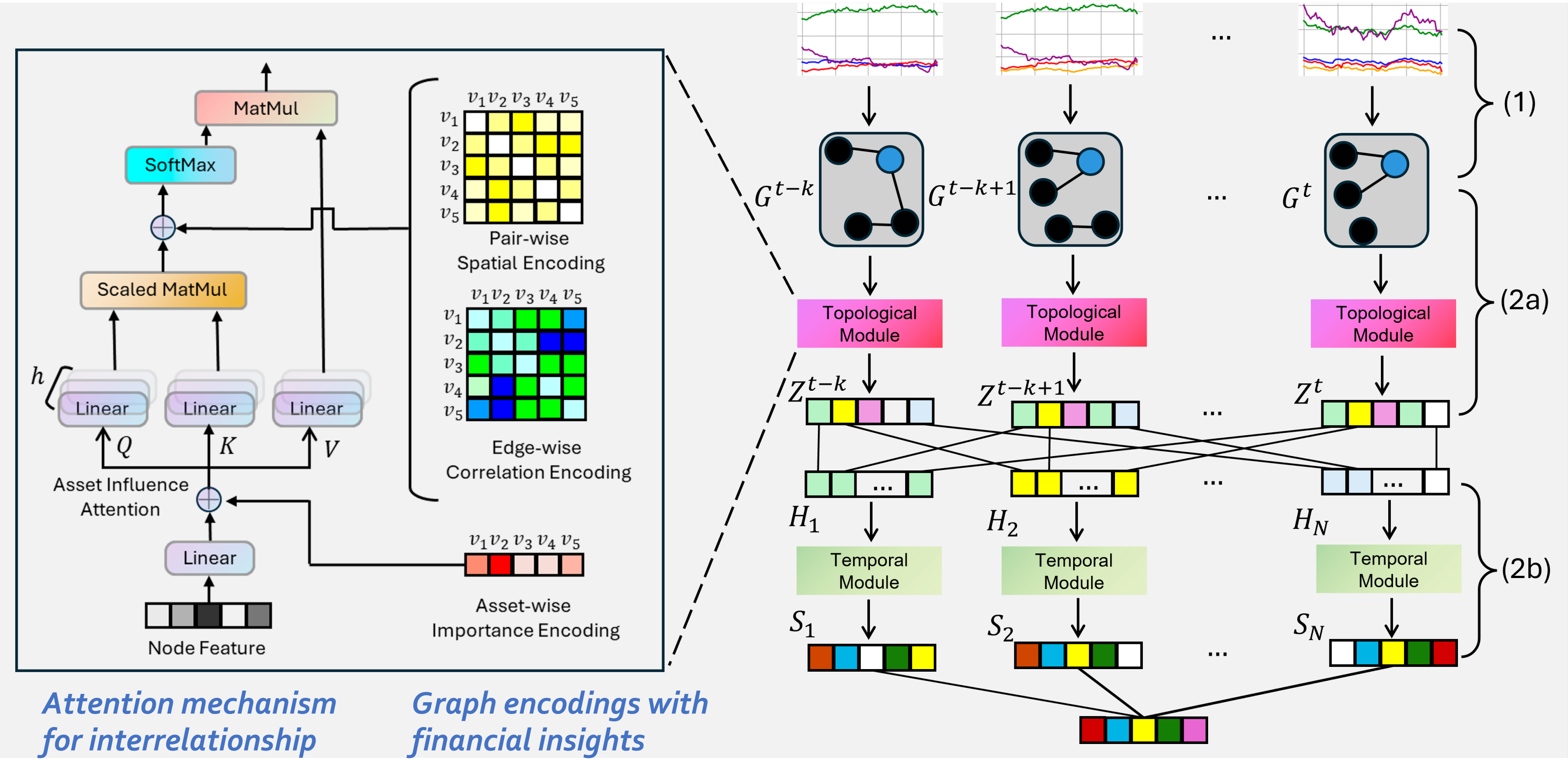


A universal formulation for dynamic graph construction in asset pricing

A graph learning model to predict the future returns of existing asset

- Dive into structural information for individual graphs
- Capture historical representations across time for individual assets

Dynamic graph representation learning framework via Spatio-Temporal Attention and Graph Encodings (DySTAGE)



A universal formulation for dynamic graph construction in asset pricing

A graph learning model to predict the future returns of existing asset

- Dive into structural information for individual graphs
- Capture historical representations across time for individual assets

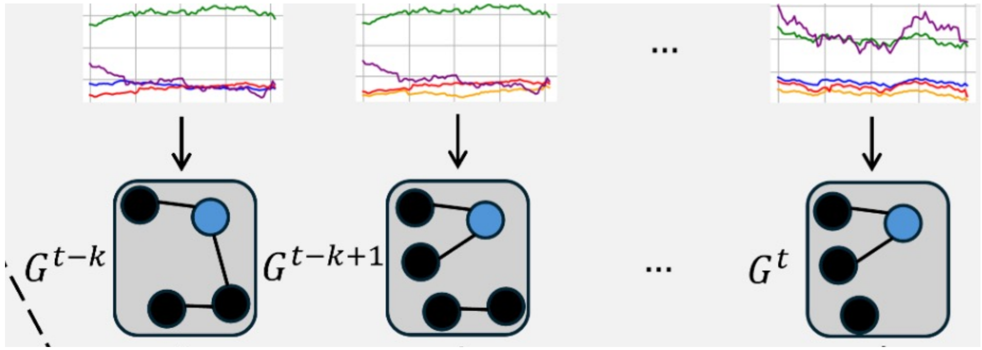
Dynamic Graph Construction from Time Series

Graph formulation

- Node: asset
- Node attributes: firm features
- Edge: strong long-term relationship

$$\mathcal{A}^t_{u,v} = \begin{cases} \rho^t_{u,v} & |\rho^t_{u,v}| > \gamma \\ 0 & |\rho^t_{u,v}| \le \gamma \end{cases}$$

- Edge attributes: multi-scale return correlations covering short-term to long-term perspectives
  - Monthly asset data: quarterly, semiannually, and yearly trends
  - Daily asset data: weekly, biweekly, monthly trends



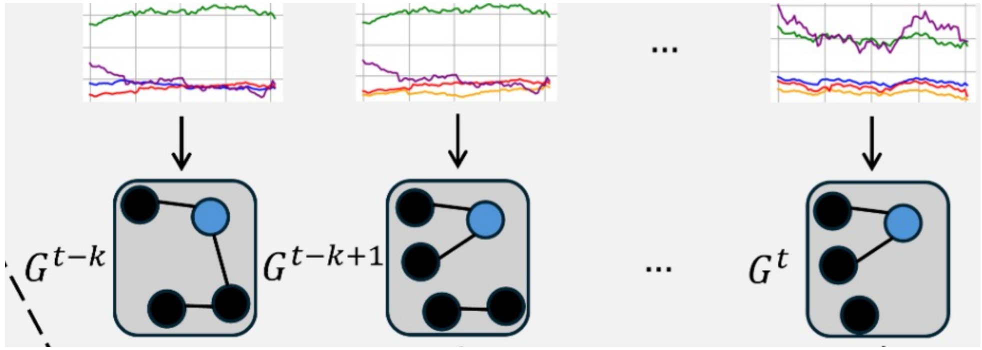
Dynamic Graph Construction from Time Series

Graph formulation

- Node: asset
- Node attributes: firm features
- Edge: strong long-term relationship

$$\mathcal{A}_{u,v}^{t} = \begin{cases} \rho_{u,v}^{t} & |\rho_{u,v}^{t}| > \gamma \\ 0 & |\rho_{u,v}^{t}| \leq \gamma \end{cases}$$

- Edge attributes: multi-scale return correlations covering short-term to long-term perspectives
  - Monthly asset data: quarterly, semiannually, and yearly trends
  - Daily asset data: weekly, biweekly, monthly trends



- Objective: given a sequence of historical graphs $\mathbb{G} = \{\mathcal{G}^{t-k}, \ldots, \mathcal{G}^{t}\}$, develop a model to predict future return for asset $y^{\{t+1\}}$:

$$\mathbf{y}^{t+1} = f(\mathcal{G}^{t-k}, \cdots, \mathcal{G}^{t}, \mathbf{X}^{t-k}, \cdots, \mathbf{X}^{t-k}, ; \theta)$$

Asset influence attention: Global interrelationship

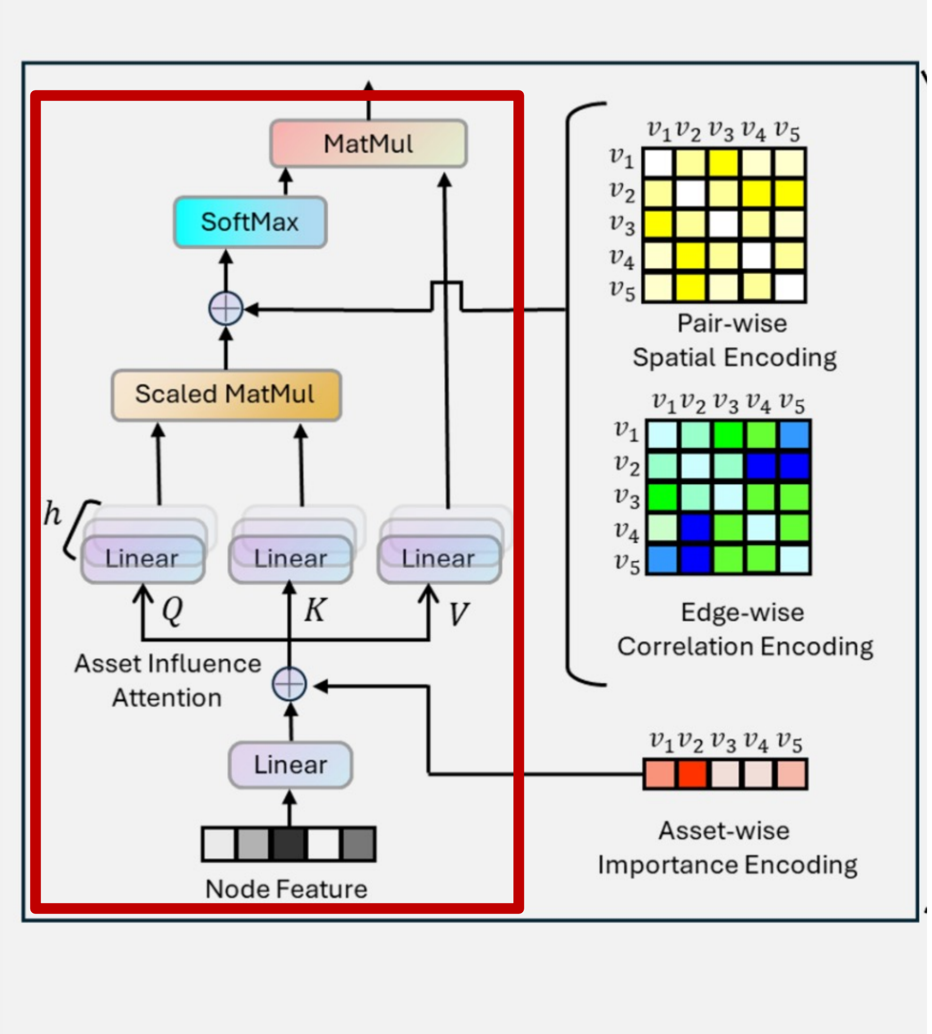
Topological Module

- Multi-head attention to capture global interrelationships between assets, non-existing assets are masked
- Layer normalization and skip connections to enhance optimization efficiency

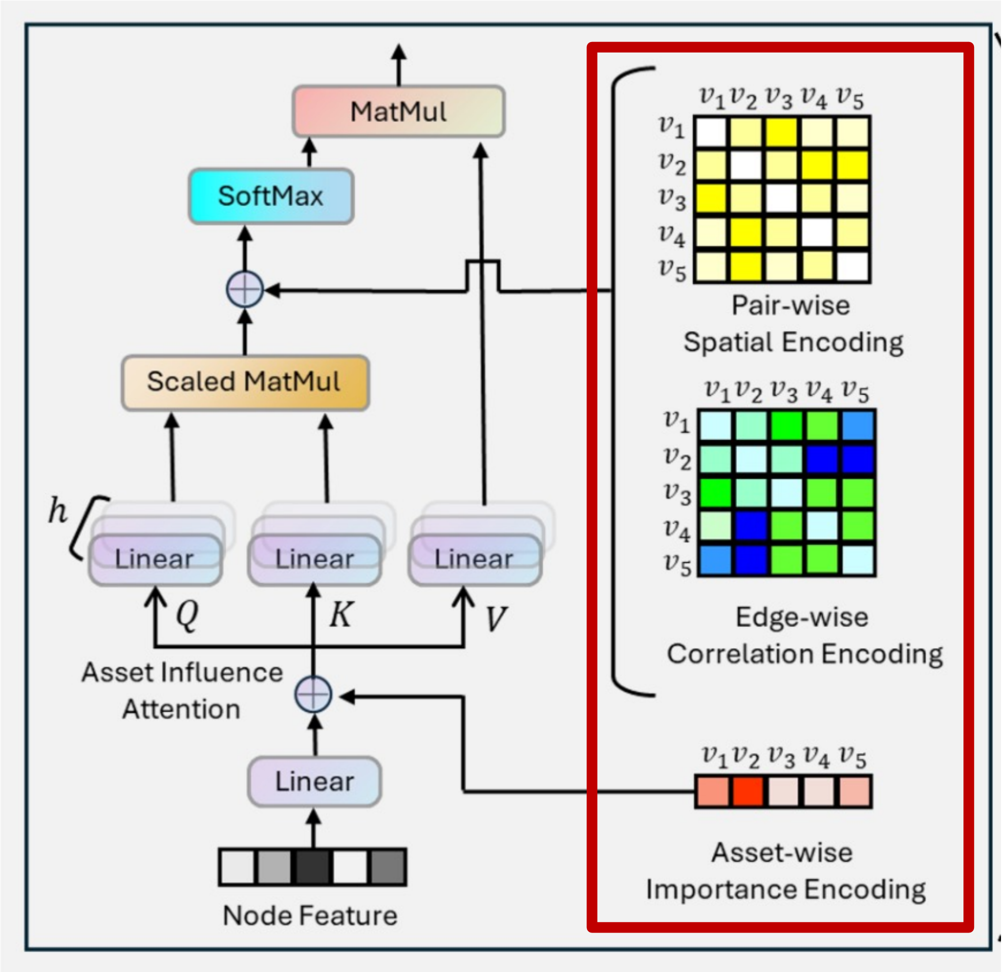
$$\mathbf{A}_h^* = \mathbf{M_0}^* \odot softmax\left(\frac{(\mathbf{X}^*\mathbf{W}_q^*)(\mathbf{X}^*\mathbf{W}_k^*)^\top}{\sqrt{d_k^*}} + \mathbf{M}_\infty{}^*\right)$$

$$\mathbf{Z} = \mathbf{X}^* + [\mathbf{A}_1^*(\mathbf{X}^*\mathbf{W}_v^*), ..., \mathbf{A}_H^*(\mathbf{X}^*\mathbf{W}_v^*)]\mathbf{W}_o^*$$

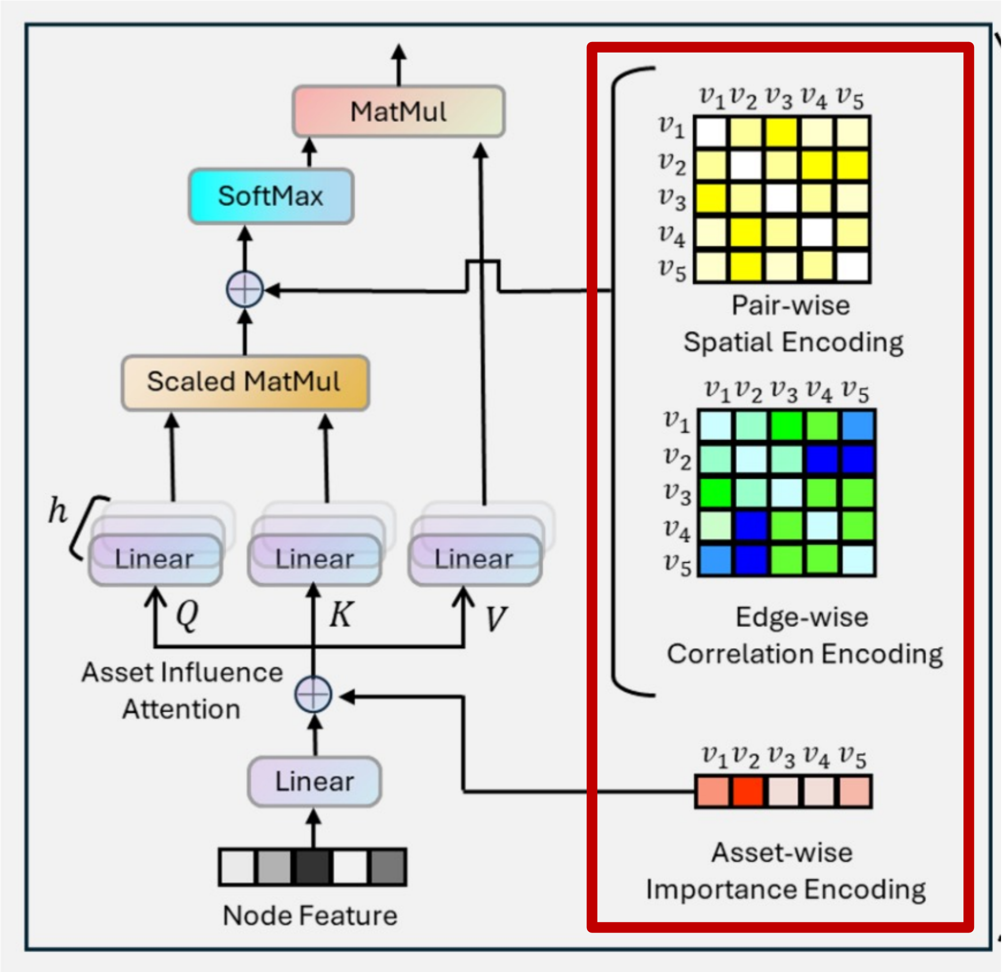
- Each element in attention matrix: influence of asset u to v
- M_∞^*: negative mask matrix
- M_0^*: zero mask matrix



Graph encodings with financial insights

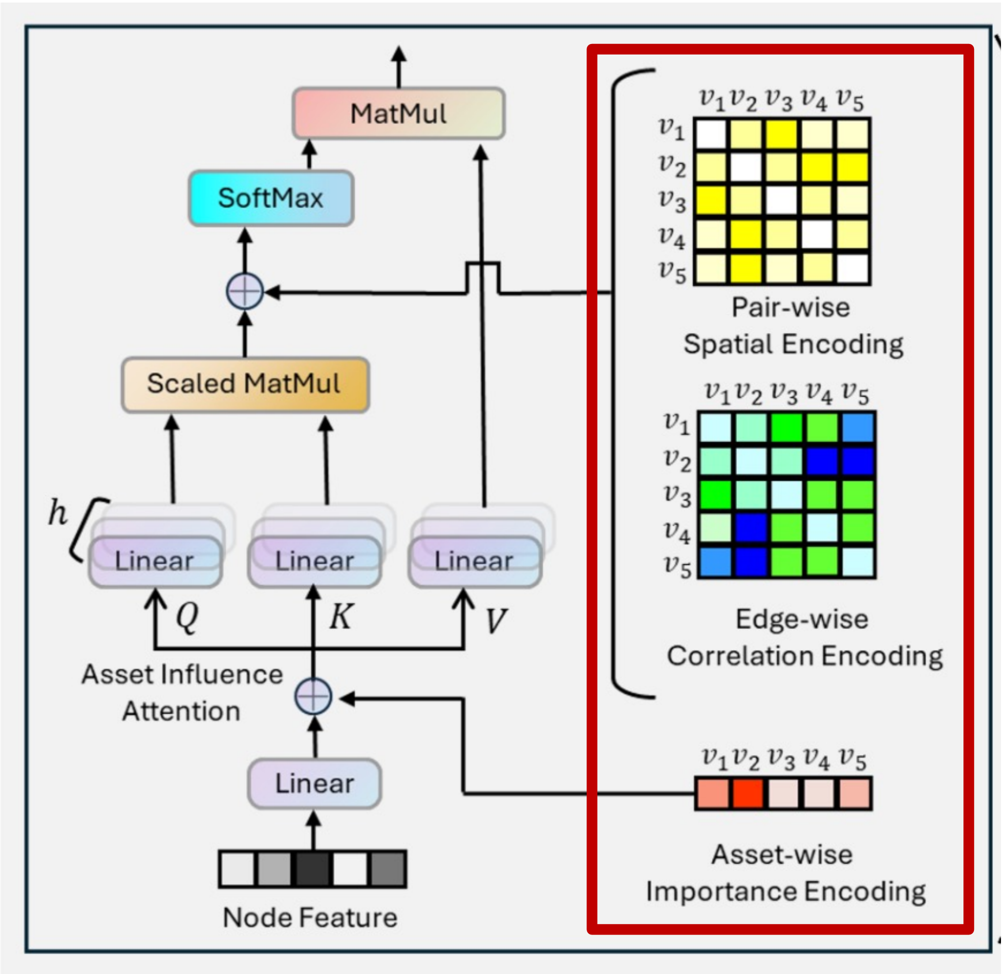


Graph encodings with financial insights



The asset with a higher node degree implies a strong correlation with a larger number of other assets, indicating its potential market impact

Graph encodings with financial insights

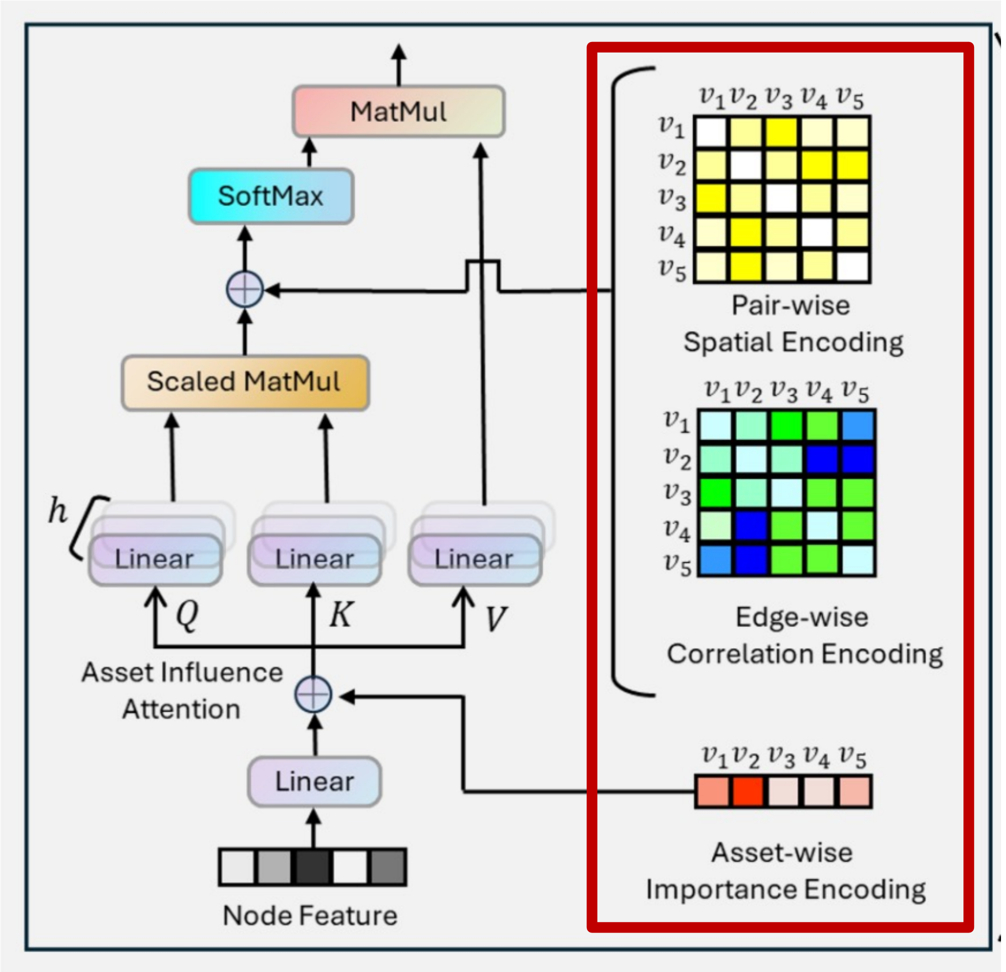


multi-scale edge attributes reveal evolving relationships over time

$$E_{u,v} = \frac{1}{p}\sum_{i}^{p}(\mathbf{c}_{u,v}\mathbf{w}_e^{\top})_i, \quad \mathbf{c}_{u,v} = \begin{cases} \mathcal{E}_{u,v} & if\, \mathcal{A}_{u,v} \neq 0 \\ \mathbf{0} & if\, \mathcal{A}_{u,v} = 0 \end{cases}$$

The asset with a higher node degree implies a strong correlation with a larger number of other assets, indicating its potential market impact

Graph encodings with financial insights



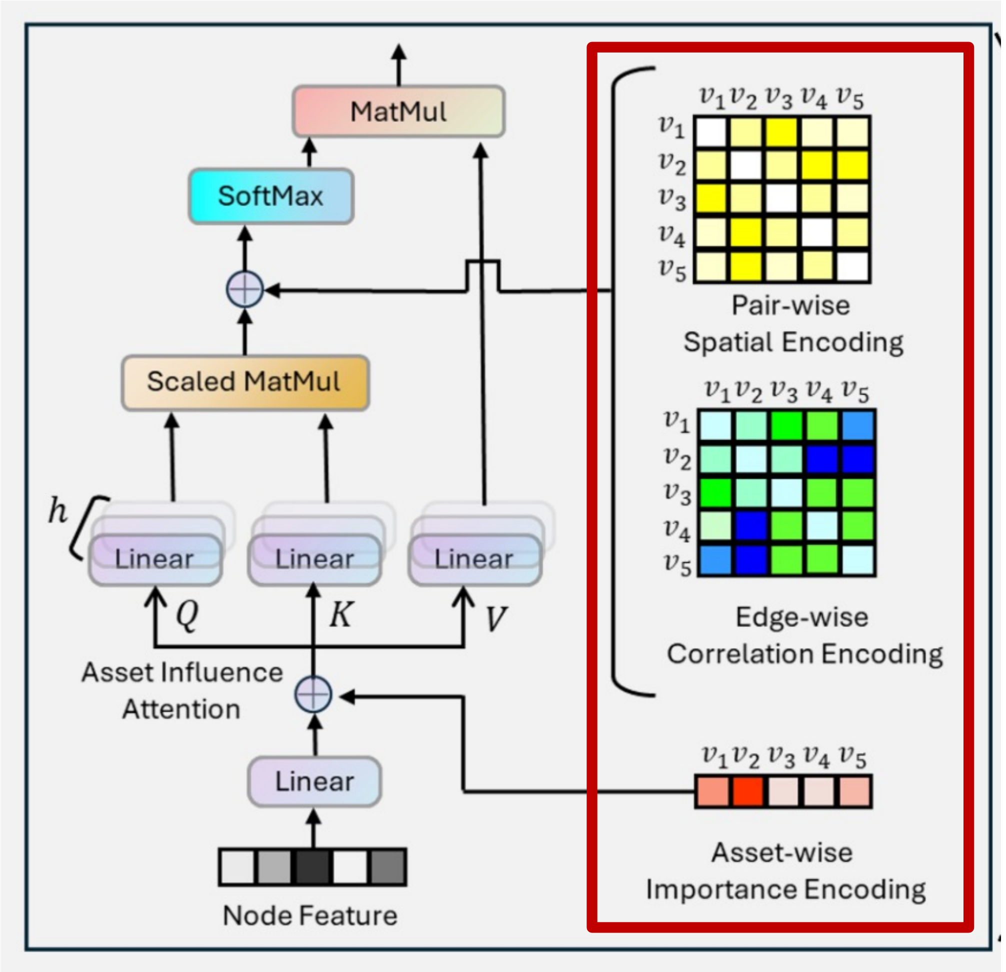
shortest path distance represents hidden connections between assets through intermediary assets even without direct correlations

multi-scale edge attributes reveal evolving relationships over time

$$E_{u,v} = \frac{1}{p} \sum_{i}^{p} (\mathbf{c}_{u,v} \mathbf{w}_e^\top)_i, \quad \mathbf{c}_{u,v} = \begin{cases} \mathcal{E}_{u,v} & if\, \mathcal{A}_{u,v} \neq 0 \\ \mathbf{0} & if\, \mathcal{A}_{u,v} = 0 \end{cases}$$

The asset with a higher node degree implies a strong correlation with a larger number of other assets, indicating its potential market impact

Graph encodings with financial insights



$$\mathbf{A}_h^* = \mathbf{M_0}^* \odot softmax\left(\frac{(\tilde{\mathbf{X}}^*\mathbf{W}_q^*)(\tilde{\mathbf{X}}^*\mathbf{W}_k^*)^\top}{\sqrt{d_k^*}} + \mathbf{S} + \mathbf{E} + \mathbf{M}_\infty{}^*\right), \quad \tilde{\mathbf{X}}^* = \mathbf{X}^* + \mathbf{D}.$$

shortest path distance represents hidden connections between assets through intermediary assets even without direct correlations

multi-scale edge attributes reveal evolving relationships over time

$$E_{u,v} = \frac{1}{p}\sum_{i}^{p}(\mathbf{c}_{u,v}\mathbf{w}_e^\top)_i, \qquad \mathbf{c}_{u,v} = \begin{cases} \mathcal{E}_{u,v} & if\,\mathcal{A}_{u,v} \neq 0 \\ \mathbf{0} & if\,\mathcal{A}_{u,v} = 0 \end{cases}$$

The asset with a higher node degree implies a strong correlation with a larger number of other assets, indicating its potential market impact

Temporal Learning

Temporal module:

- investigate historical representations along the temporal dimension for each node individually

$$\mathbf{A}^{\star}_{h'} = softmax(\frac{(\tilde{\mathbf{H}}\mathbf{W}^{\star}_{q})(\tilde{\mathbf{H}}\mathbf{W}^{\star}_{k})^{\top}}{\sqrt{d^{\star}_{k}}} + \mathbf{M}_{\infty}{}^{\star})$$

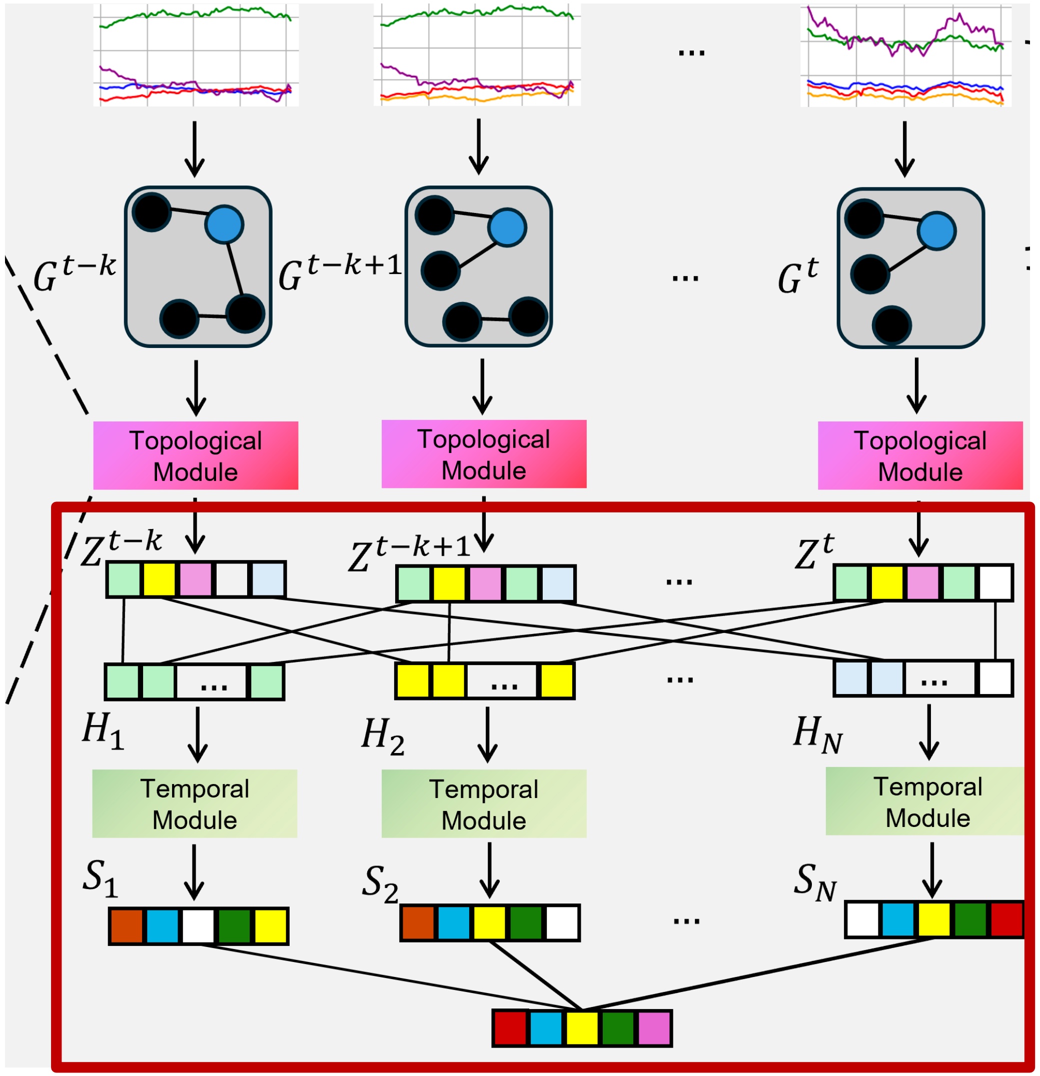
$$\mathbf{S} = \mathbf{H} + [\mathbf{A}^{\star}_{1}(\tilde{\mathbf{H}}\mathbf{W}^{\star}_{v}), \dots, \mathbf{A}^{\star}_{H}(\mathbf{H}\mathbf{W}^{\star}_{v})]\mathbf{W}^{\star}_{o}$$

- Node-level Prediction

$$\hat{y}_u = M^{(P)} \cdot \tilde{y}_u, \quad \tilde{y}_u = tanh(MLP(\mathbf{S_u}))$$

$$\mathcal{L} = \sum_{u \in \mathcal{V}} (\hat{y}_u - y_u)^2$$

  - $\hat{y}_u$: predicted return for asset u at the time step $t+1$



Research Questions and Datasets

- Performance compared with other dynamic and static methods
- Investment advice and profitability in real-world scenarios
- Graph representation learning on asset patterns
- Contribution of each component in DySTAGE

Research Questions and Datasets

- Performance compared with other dynamic and static methods
- Investment advice and profitability in real-world scenarios
- Graph representation learning on asset patterns
- Contribution of each component in DySTAGE

Table 1: Summary of Statistics for Dynamic Asset Graph Datasets. Each entry details the snapshot counts, total nodes appearing across the entire time range, average number of edges per snapshot, feature counts, time lag, horizon, data collection frequency, and time span.

Dataset	# Snapshots	# Nodes	# Edges	# Features	Lag	Horizon	Frequency	Range
Russell 3000	193	2,151	1,290K	166	12	1	Monthly	Jan 2000 - Dec 2021
MLFI	172	990	387K	93	12	1	Monthly	Jan 2000 - Mar 2019
S&P 500	2,396	460	123K	24	20	1	Daily	Jan 3, 2011 - Dec 31, 2020

Performance on Asset Pricing (RQ1)

Table 3: Comparison results from benchmarks and our model. MAPE resuls are in the form of percentage (%).

Type	Model	Russell 3000			MLFI			S&P 500		
		RMSE	MAE	MAPE	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Time Series	ARIMA	0.1798	0.1422	117.8364	0.1254	0.0970	87.0359	0.0223	0.0178	17.8370
	N-Beats	0.1327	0.1050	83.8827	0.1005	0.0793	69.7604	0.0195	0.0158	15.8619
Static GNN	GAT	0.1073	0.0874	66.3047	0.0802	0.0651	55.6068	0.0154	0.0131	13.1286
	GraphSAGE	0.1060	0.0864	65.4026	0.0811	0.0656	56.0802	0.0156	0.0131	13.1464
	ARMAConv	0.1081	0.0877	66.2636	0.0808	0.0653	55.8012	0.0158	0.0133	13.2557
	UniMP	0.1078	0.0853	64.1338	0.1078	0.0853	64.1331	0.0156	0.0134	13.3536
Dynamic GNN	DySAT	0.1039	0.0840	63.2542	0.0806	0.0652	55.7200	0.0155	0.0132	13.2118
	DY-GAP	0.1357	0.1089	87.7335	0.0806	0.0652	55.7378	0.0155	0.0131	13.0723
	T-GCN	0.1078	0.0882	67.0031	0.0813	0.0657	56.0604	0.0168	0.0145	14.5104
	EvolveGCN	0.1064	0.0845	63.5845	0.0806	0.0651	55.5625	0.0155	0.0132	13.1989
	GCLSTM	0.1033	0.0839	62.9582	0.0807	0.0650	55.5467	0.0156	0.0134	13.4025
	DyTed	0.1040	0.0844	63.3368	0.0800	0.0647	55.3379	0.0155	0.0132	13.1666
	DGIB	0.1031	0.0837	62.8114	0.0802	0.0649	55.4483	0.0154	0.0131	13.0751
	DySTAGE	**0.1026**	**0.0833**	**62.5027**	**0.0797**	**0.0644**	**54.9632**	**0.0154**	**0.0131**	**13.0602**

Portfolio Management (RQ2)

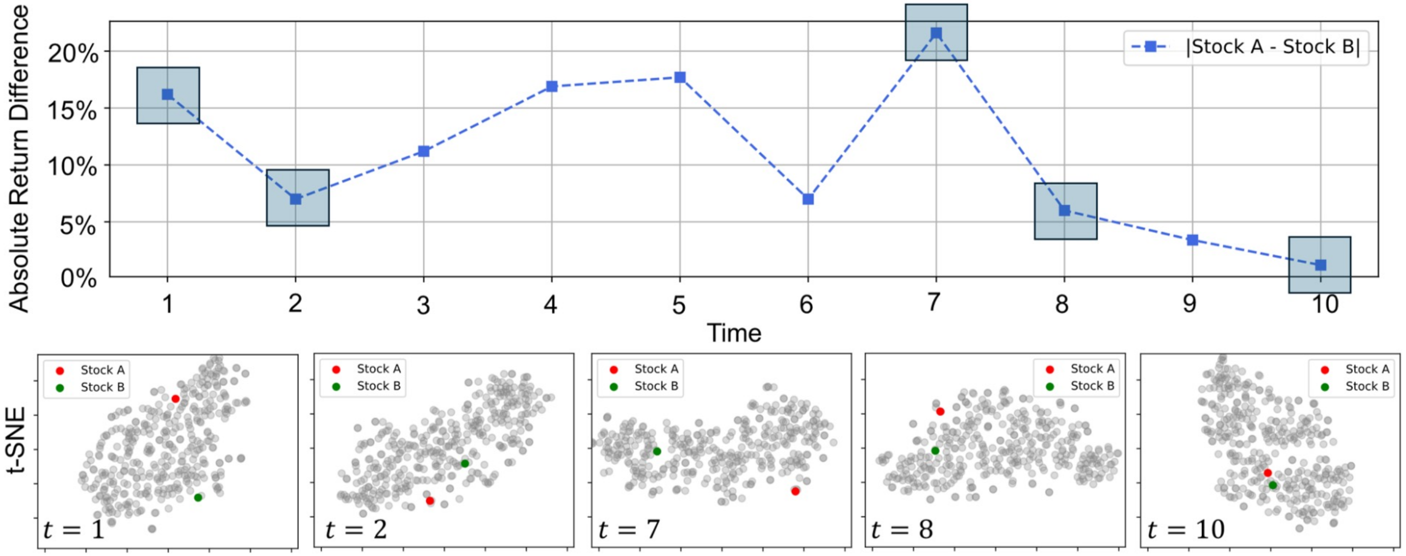
Invest in long positions on assets with top 10% highest predicted excess returns and assign them equal weight

Table 4: Portfolio management results on the three datasets. CR and AR are in the format of percentage (%). ↑ means the larger the better.

Type	Model	Russell 3000			MLFI			S&P 500		
		CR(%) ↑	AR(%) ↑	SR ↑	CR(%) ↑	AR(%) ↑	SR ↑	CR(%) ↑	AR(%) ↑	SR ↑
Time Series	ARIMA	42.1388	20.1368	1.0047	0.9074	0.5435	0.1198	17.6767	12.1484	0.7275
	N-Beats	49.5191	23.3519	1.1667	9.2134	5.0084	0.4069	25.9006	18.0936	1.0282
Static GNN	GAT	42.7684	20.4142	1.1837	8.0401	4.7492	0.3549	10.5096	7.4825	0.5219
	GraphSAGE	49.8937	23.5131	**1.2077**	-0.7568	-0.4547	0.0483	24.5304	17.1641	0.9445
	ARMAConv	25.7004	12.6751	0.7184	-0.5407	-0.3247	0.0675	14.1315	10.0146	0.6152
	UniMP	40.0290	19.2031	1.0231	5.8182	3.4514	0.2748	15.6653	1.0802	0.7083
Dynamic GNN	DySAT	37.9606	18.2812	1.0156	-0.3293	-0.1977	0.0808	23.3353	16.3512	0.9530
	DY-GAP	40.8272	19.5572	1.1609	9.8075	5.7740	0.4285	19.0134	13.3927	0.8330
	T-GCN	35.7086	17.2698	0.9543	7.6595	4.5486	0.3626	6.0337	4.3211	0.3410
	EvolveGCN	29.7708	14.5642	0.8721	2.2540	1.3464	0.1611	5.1677	3.7051	0.2837
	GCLSTM	41.9357	20.4726	1.0616	-3.2744	-1.9777	-0.0330	8.0919	5.7793	0.4216
	DyTed	40.1514	19.2575	0.9667	-0.7692	-0.4622	0.0643	9.4853	6.5678	0.5208
	DGIB	29.2723	14.3344	0.9369	6.1102	3.6226	0.3047	8.5284	6.0876	0.4149
	DySTAGE	**50.3428**	**23.7152**	1.1975	**10.2829**	**6.0486**	**0.4614**	**31.5506**	**21.8969**	**1.2945**

- DySTAGE consistently generates the highest return with strong balance between profitability and risk management, offers lucrative investment recommendations in real-world scenarios

Graph Learning (RQ3)



Spatial distribution in embeddings effectively mirrors the actual financial performance disparities.

Ablation Study (RQ4)

Model	Russell	MLFI	S&P
w/o Importance	62.7537	55.3018	13.0674
w/o Temporal	62.6115	55.0411	13.1801
w/o Spatial	62.5868	54.8906	13.0734
w/o Edge	62.5943	54.9654	13.0690
DySTAGE	**62.5027**	**54.9632**	**13.0602**

- Temporal module significantly boosts the performance, while DySTAGE equipped solely with the topological module is remarkably powerful
- All graph encodings contribute to the model improvement. Asset-wise Importance encoding is the most influential component

Conclusion

We introduce DySTAGE, a novel dynamic graph representation learning framework for asset pricing.

DySTAGE effectively captures both topological and temporal patterns, utilizing graph encodings within the financial network.

Extensive experiments proves the superiority of DySTAGE over conventional and popular benchmarks in predictive accuracy

Thank you! Questions?

LinkedIn

Homepage