Systemic Risk Modeling in Listed Chinese Financial Institutions: A Two-Layer Network Analysis with a Graph Attention Prototype

Zelin Guo[1], Xinrui Wang[2]
1. Shanghai University of International Business and Economics, Shanghai, China
2. Inner Mongolia Agricultural University, Hohhot, Inner Mongolia, China
DOI: 10.65285/HSS.V2I4.001

Abstract: System-wide financial risk may be obscured when return comovement and directional spillovers are examined separately. We analyzed 3,167 daily observations for 24 listed Chinese banks, securities firms, and insurers from 19 December 2011 to 31 December 2024. A two-layer multiplex network combined a thresholded Pearson return-comovement layer with a LASSO–VAR generalized forecast error variance decomposition layer, and maximum spanning arborescences represented dominant transmission backbones. Across four non-overlapping windows, verified total connectedness was 88.50%, 89.01%, 89.37%, and 86.34%, whereas return-comovement density was 0.736, 0.739, 0.833, and 0.518. Full-sample estimates identified GDB, JTB, HtS, ZXS, and ZSB as leading net transmitters and GJS and GFS as leading receivers. Correcting the archived edge direction and arborescence implementation produced period-specific roots rather than a stable single source. The accompanying GATv2 prototype merged the two relations as edge features, but only four network snapshots, circular structural targets, absent period-5 inputs, and no reproducible temporal split prevented verification of the reported clustering and classification scores. The graph-attention evidence is therefore treated as a proof-of-concept, not out-of-sample early warning tool. This audited framework separates synchronous returns from directional connectedness and may support transparent monitoring after prospective validation.
Key words: systemic financial risk; return spillovers; multiplex financial network; LASSO–VAR; graph attention network; connectedness

1. Introduction

Systemic financial risk is a property of interdependence rather than a simple sum of institution-level risks. Network links can diversify exposures in normal conditions yet transmit losses or synchronized behavior under stress [1,2]. Market-based connectedness measures are therefore useful for identifying changes in joint behavior, but correlation alone does not reveal the direction of transmission and should not be interpreted as causality.

Forecast error variance decompositions provide a directional view of connectedness. Diebold and Yilmaz introduced return and volatility spillover indices and later linked variance decompositions to weighted directed networks [3–5]. In high-dimensional systems, LASSO regularization can reduce the number of VAR coefficients and make rolling connectedness estimation more tractable [6–9]. These methods quantify statistical transmission, not default contagion, and their interpretation depends on lag order, forecast horizon, regularization, and window length.

Multiplex representations retain a common node set while distinguishing edge layers [10–14]. Here, return comovement and directional forecast-error contributions encode different relations. Community detection summarizes mesoscale structure [15], while optimum branchings provide a parsimonious backbone without recovering a complete economic mechanism [16].

Graph attention networks learn neighbor-specific aggregation weights [17], and GATv2 makes attention rankings conditional on the query node [18]. Risk identification nevertheless requires independent labels, temporal splits, adequate graph samples, and baselines; otherwise, model-assigned states are structural summaries rather than verified crises.

We compare both layers across four windows, identify transmitters, receivers, intermediaries, and corrected dominant backbones, and audit whether graph attention supports early-warning claims. Verified network findings are separated from the unvalidated GAT prototype.

2. Materials and Methods

2.1 Data and preprocessing

The processed panel contained 3,167 trading days for 24 listed institutions: 14 banks, six securities firms, and four insurers, from 19 December 2011 to 31 December 2024. The archived manuscript described closing prices and market-model imputation against the CSI 300. The final panel had no missing values, but the initial missing count, benchmark file, and imputation code were absent. Daily log return was

$$r(i,t) = \ln\left[\frac{P(i,t)}{P(i,t-1)}\right]$$

where $P(i,t)$ is the closing price of institution i on day t. Means in the source descriptive table were reported in percentage units; for example, 0.039 denotes 0.039%, not 3.9%. Component output supported stationarity of all return series at conventional critical values, although several ADF entries in the final Chinese manuscript had been copied incorrectly and were replaced by the component-report values in the Supplementary Materials.

Four non-overlapping windows were used: 19 December 2011–27 March 2015; 30 March 2015–22 June 2018; 25 June 2018–17 September 2021; and 22 September 2021–31 December 2024. They contained 791, 791, 791, and 794 observations, respectively.

Table 1. Data and verified model settings.

Item	Verified setting
Sample	24 institutions; 14 banks, 6 securities firms, 4 insurers
Processed observations	3,167 daily returns; 19 December 2011–31 December 2024
Return layer	Pearson correlation; edge when absolute correlation exceeds 0.5
Spillover layer	Row-normalized GFEVD; edge j -> i for contribution of shock j to i
Rolling windows	120 days; 180 and 250 days reported as sensitivity settings
Audited GAT code	Two GATv2 layers; hidden 16; heads 2/1; Adam; learning rate 0.01; 300 epochs; MSE

2.2 LASSO–VAR spillover estimation

For an N-dimensional return vector r(t), the source methodology specified

$$r(t) = c + \sum_{l=1}^{p} A(l)\, r(t-1) + \varepsilon(t)$$

LASSO estimation minimized the equation-level squared error plus an L1 penalty,

$$\hat{\beta}_i = \arg\min_{\beta_i} \left\{ \sum_t \left[r(i,t) - x(t)^T \beta_i \right]^2 + \lambda \|\beta_i\|_1 \right\}$$

The H-step generalized forecast error variance contribution from shock j to variable i was denoted $\theta_{ij}(H)$ and row-normalized as

$$\tilde{\theta}_{ij}(H)=\frac{\theta_{ij}(H)}{\sum_{j=1}^{N}\theta_{ij}(H)}$$

Thus, $FROM_i=100\sum_{j\neq i}\tilde{\theta}_{ij}(H)$, $TO_i=100\sum_{j\neq i}\tilde{\theta}_{ji}(H)$, $NET_i=TO_i-FROM_i$, and total connectedness was $TCI=\frac{100}{N}\sum_{i\neq j}\tilde{\theta}_{ij}(H)$. The archived dynamic table reversed these labels and the net sign; both were corrected.

The primary rolling window was reported as 120 trading days, with 180- and 250-day sensitivity plots. VAR lag order, selection criterion, lambda, H, scaling, and the LASSO–VAR program were not archived and remain [parameters to be confirmed]. Consequently, the four supplied spillover matrices were audited and recomputed descriptively, but the rolling LASSO–VAR series could not be regenerated.

2.3 Construction of the two-layer financial network

The node-aligned multiplex network contained the same 24 institutions in both layers (Fig. 1). Layer 1 was an undirected return-comovement network. An edge joined i and j when $|\rho_{ij}|>0.5$, with Pearson correlation as weight. No negative correlation exceeded the threshold in any window. Density, absolute weighted strength, PageRank, and Louvain modularity summarized this layer.

Layer 2 was a directed return-spillover network based on $\tilde{\cdot}(H)$. Because row i, column j measures the contribution of shock j to i, the directed edge was j > i, weighted by $100\tilde{\theta}_{ij}(H)$. Full positive matrices were used for strengths and connectedness. The source visualization used a 25th-percentile filter and retained at most 100 edges; that display rule was not used for quantitative inference.

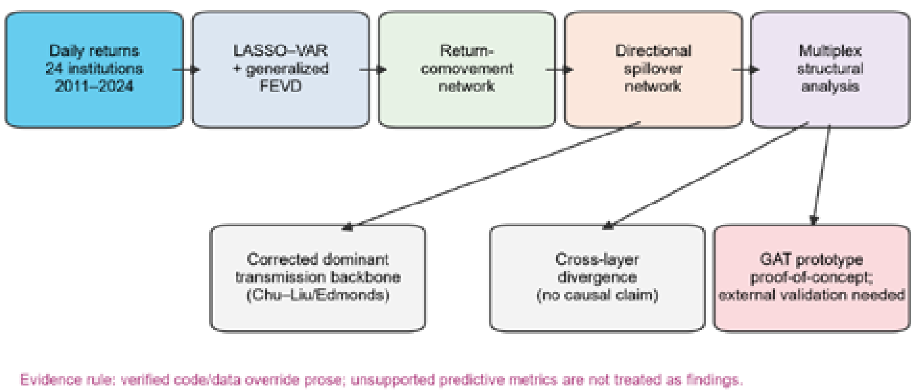

Figure 1. Evidence-aware analytical framework. Verified returns and supplied spillover matrices support the two-layer analysis. The graph-attention component is separated as a prototype requiring external validation.

2.4 Identification of dominant risk-transmission paths

For each directed spillover layer, the ChuLiuEdmonds algorithm selected the maximum-total-weight spanning arborescence [16]. We retained positive spillover weights and used the economically consistent j > i direction. No sign reversal was applied. The optimized arborescence determined its root as the unique node with zero in-degree; betweenness did not preselect the root. This corrects the source script, which negated weights before maximizing and calculated a putative root without enforcing it. The resulting tree is interpreted only as a strongest-path representation. Weighted betweenness used inverse spillover weight as path length to identify potential intermediaries in the full network.

2.5 Graph attention-based risk identification

Two incompatible prototypes were archived. The period-5 script used GATv2 with mean return and standard deviation as node features and correlation/spillover as edge attributes. It had two layers, hidden dimension 16, heads 2 then 1, ReLU, Adam (0.01), 300 epochs, and mean-squared error. It regressed weighted degree, betweenness, and net spillover on four graphs; dropout, early stopping, a seed, class balancing, and data partitions were absent.

The relation-typed script referenced missing filenames and undefined metadata, supplied no labels, and used an incompatible model call. Its feature files also included centralities, strengths, net spillover, and cross-layer distance, creating circularity with structure-driven labels. Neither script implemented the claimed GRU, eight heads, hidden dimension 256, or 500 epochs. We therefore use **GATv2 prototype**, not HGAT, and reject a validated-forecasting interpretation.

2.6 Evaluation metrics

The source reported silhouette 0.567, Calinski–Harabasz 1478.097, study-specific risk stability 0.003, and F1 0.775. Stability was described as the class-mean standard deviation of predicted probabilities. No executable calculation, labels, F1 convention, probability file, seed, or temporal split was found; period-5 inputs were absent. These values and the 7/10/7 risk counts remain audit records, not findings.

3. Results

3.1 Static and dynamic spillover patterns

The full-sample table gave positive net spillovers for 13 institutions. The leading transmitters were China Everbright Bank (GDB, 15.85 percentage points), Bank of Communications (JTB, 9.50), Huatai Securities (HtS, 7.77), CITIC Securities (ZXS, 7.36), and China Merchants Bank (ZSB, 7.06). Guojin Securities (GJS, −27.81) and GF Securities (GFS, −15.33) were the largest receivers. A high FROM value means exposure to others' shocks; it does not by itself indicate resilience.

Window-level total connectedness remained high: 88.50%, 89.01%, 89.37%, and 86.34% (Fig. 2; Table 2). The source-only 120-day curve ranged approximately from 70% to 93%, but the series and 180-/250-day outputs were unavailable; peaks and robustness were not independently quantified.

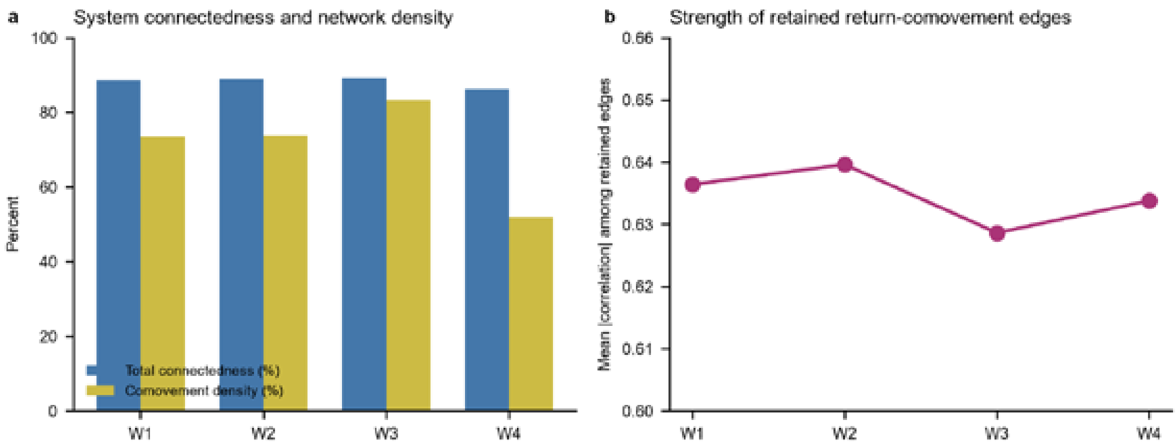

Figure 2. Window-level connectedness and return-comovement structure. (a) Total connectedness from supplied GFEVD matrices and density of the thresholded return network. (b) Mean absolute correlation among retained edges.

Table 2. Verified network results across four windows.

Window	Density	Mean absolute correlation	Total connectedness (%)	Louvain modularity	Corrected backbone root
2011–2015	0.7355	0.6364	88.50	0.1217	PFB
2015–2018	0.7391	0.6396	89.01	0.1112	XYB
2018–2021	0.8333	0.6286	89.37	0.0933	HXB
2021–2024	0.5181	0.6338	86.34	0.3209	PAI

3.2 Evolution of the return-comovement network

Comovement density was 0.736, 0.739, 0.833, and 0.518, whereas mean retained-edge weight was 0.636, 0.640, 0.629, and 0.634. Thus, Window 3 added many moderately strong edges, while Window 4 contracted sharply without materially weakening the average retained edge (Fig. 3). The strongest nodes by absolute weighted degree changed from PFB, GDB, ZSB, JTB, and NBB in Window 1 to PAI, PFB, TBI, XHI, and PAB in Window 4.

Louvain modularity was 0.1217, 0.1112, 0.0933, and 0.3209; the number of communities was two, three, two, and two. The final window therefore combined lower density with more pronounced modular separation. These are descriptive structures conditional on the fixed 0.5 threshold.

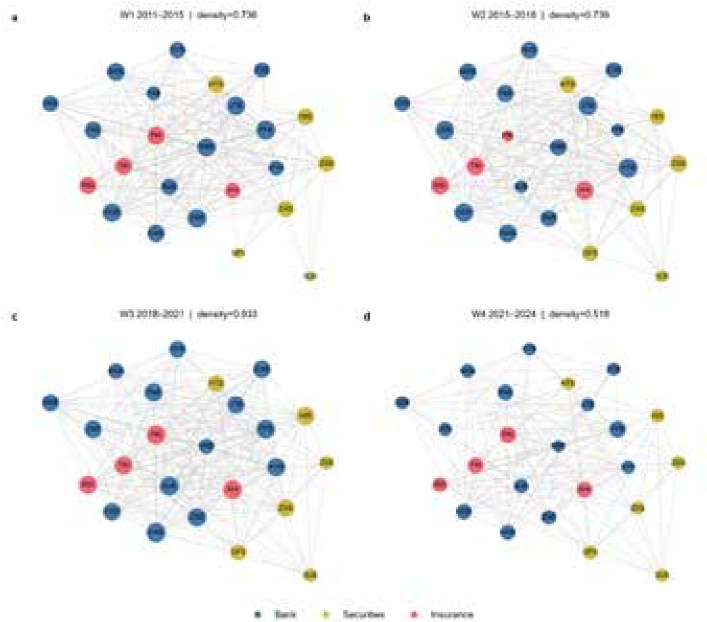

Figure 3. Evolution of the return-comovement network. Node color denotes sector and node size denotes absolute weighted strength. Positions are fixed across panels.

3.3 Risk transmitters, receivers, and dominant transmission paths

After correcting direction, HtS and TBI were net transmitters in all four windows, whereas GSB and GJS were receivers in all four. Other roles were episodic. PFB led Window 1 (22.77), XYB led Window 2 (24.09), HXB led Window 3 (16.44), and PAI led Window 4 (18.35). GJS was the largest receiver in Windows 1 and 3 (−39.35 and −18.44), while PAI and HTS were the largest receivers in Windows 2 and 4 (−28.97 and −13.73).

The corrected maximum arborescences had roots PFB, XYB, HXB, and PAI. Full-network inverse-weight betweenness highlighted HtS in Windows 1 and 3, XYB/ZSS in Window 2, and PFB in Window 4. The changing roots and intermediaries argue against a time-invariant propagation source (Fig. 4).

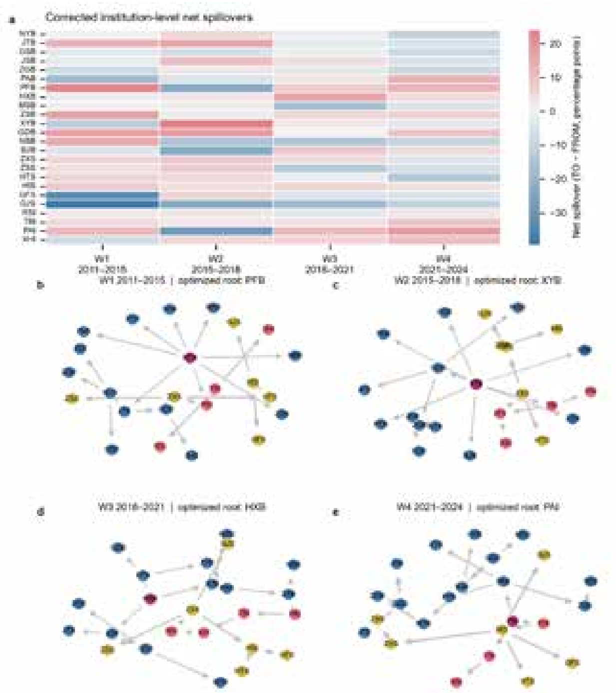

Figure 4. Corrected net spillovers and dominant transmission backbones. (a) Net spillover is TO minus FROM. (b–e) Maximum spanning arborescences using positive directed weights; purple nodes are optimization-derived roots.

3.4 Cross-layer structural coupling

The archived cross-layer statistic was the Euclidean distance between each institution's raw correlation row and raw spillover row, followed by min–max scaling across institutions within each window. Recalculation reproduced the supplied values to numerical precision. Mean normalized distance was 0.462, 0.355, 0.412, and 0.431. The most divergent nodes were ZGB in Window 1, GJS in Windows 2 and 3, and NYB in Window 4 (Fig. 5).

This metric measures row-profile divergence, not causal coupling. Because the two rows use different scales and normalization is repeated within each window, absolute comparisons over time are limited. It is best interpreted as a screening statistic for institutions whose comovement and directional-spillover profiles are relatively misaligned.

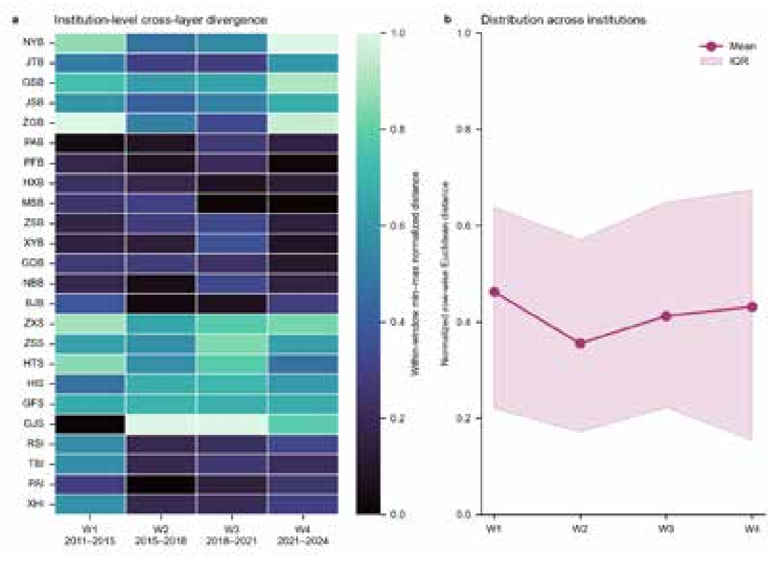

Figure 5. Cross-layer structural divergence. (a) Within-window min–max normalized row-wise Euclidean distance. (b) Mean and interquartile range across institutions.

3.5 GAT-based risk identification and early-warning results

No GAT performance result met the verification standard. The scores appeared only in a manuscript table, without probabilities or a reproducible classifier. The period-5 script required an absent adjacency matrix and 2025–2026 returns from data ending in 2024; it also regressed contemporaneous network statistics rather than crisis labels. Period-5 states and paths are therefore excluded.

Table 3. Audit status of reported GAT results.

Reported item	Source value	Verification status
Silhouette coefficient	0.567	No labels or executable calculation
Calinski–Harabasz index	1478.097	No clustering output or calculation
Study-specific risk stability	0.003	Definition described; probability file and code absent
F1 score	0.775	Averaging convention, split, labels, and code absent
Period-5 risk counts	7/10/7	Period-5 adjacency and return data absent

4. Discussion

Correlation networks describe synchronous prices, whereas GFEVD networks allocate directional forecast-error contributions. In Window 4, comovement density fell to 0.518 while directional connectedness remained 86.34%; sparse synchronous links can coexist with broadly distributed directional dependence.

Transmitter and receiver roles changed with topology [2,10–12]. HtS and TBI were persistent transmitters, GSB and GJS persistent receivers, and other institutions switched roles. Sector labels cannot establish resilience. Balance sheets, trading, common exposures, policy support, and microstructure are plausible but unobserved mechanisms.

The corrected arborescences are surveillance maps, not causal mechanisms. Roots and high-betweenness intermediaries may prioritize examination, but intervention requires exposure, liquidity, capitalization, and stress-test evidence.

With four snapshots, no external labels or temporal holdout, and structural outcomes among candidate features, the prototypes cannot establish forecasting. GATv2 remains plausible for combining edge attributes but requires independent labels, prospective tests, and simpler baselines.

Several limitations define the scope. The sample covers only 24 listed institutions in China and excludes funds, trusts, smaller firms, and unlisted institutions. The return layer depends on a fixed correlation threshold; no threshold sensitivity analysis was archived. LASSO–VAR results depend on lag order, penalty selection, forecast horizon, and rolling-window length, several of which could not be verified. Spillover edges represent statistical dependence rather than structural causation. The cross-layer distance mixes differently scaled rows and is normalized separately by window. The GAT sample is extremely small, risk states are structure-driven proxies, and neither external outcomes nor multimodel baselines were available. Results should therefore not be generalized beyond the sampled market or presented as deployment-ready early warning.

5. Conclusion

The two-layer network revealed that return comovement and directional return spillovers evolved differently among 24 listed Chinese financial institutions. Directional connectedness remained above 86% in all four windows, whereas the return-comovement network expanded in 2018–2021 and contracted markedly in 2021–2024. Institution roles were heterogeneous: HtS and TBI were recurrent transmitters, GSB and GJS recurrent receivers, and several banks, securities firms, and insurers switched roles across periods. Corrected Chu–Liu/Edmonds arborescences identified different roots and intermediaries in each window, providing a parsimonious map of dominant statistical transmission paths rather than a causal contagion model. The archived GATv2 prototype did not support a verified forecasting claim because its targets and candidate inputs were structurally circular, only four snapshots were available, period-5 inputs were absent, and no reproducible temporal evaluation existed. For macroprudential monitoring, the network results support tracking changing transmitters, receivers, and structural backbones, but model-based early warning requires independently defined crisis labels, prospective temporal validation, and comparison with transparent baselines. The principal limitation is therefore not graph-attention capacity but the current evidence design and reproducibility of the econometric pipeline.

References

[1] Haldane AG, May RM. Systemic risk in banking ecosystems. Nature. 2011;469:351–355.

[2] Billio M, Getmansky M, Lo AW, Pelizzon L. Econometric measures of connectedness and systemic risk in the finance and insurance sectors. Journal of Financial Economics. 2012;104:535–559.

[3] Diebold FX, Yilmaz K. Measuring financial asset return and volatility spillovers, with application to global equity markets. Economic Journal. 2009;119:158–171.

[4] Diebold FX, Yilmaz K. Better to give than to receive: predictive directional measurement of volatility spillovers. International Journal of Forecasting. 2012;28:57–66.

[5] Diebold FX, Yilmaz K. On the network topology of variance decompositions: measuring the connectedness of financial firms. Journal of Econometrics. 2014;182:119–134.

[6] Demirer M, Diebold FX, Liu L, Yilmaz K. Estimating global bank network connectedness. Journal of Applied Econometrics. 2018;33:1–15.

[7] Koop G, Pesaran MH, Potter SM. Impulse response analysis in nonlinear multivariate models. Journal of Econometrics. 1996;74:119–147.

[8] Pesaran HH, Shin Y. Generalized impulse response analysis in linear multivariate models. Economics Letters. 1998;58:17–29.

[9] Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B. 1996;58:267–288.

[10] Caccioli F, Barucca P, Kobayashi T. Network models of financial systemic risk: a review. Journal of Computational Social Science. 2018;1:81–114.

[11] Battiston S, Puliga M, Kaushik R, Tasca P, Caldarelli G. DebtRank: too central to fail? Financial networks, the FED and systemic risk. Scientific Reports. 2012;2:541.

[12] Poledna S, Molina-Borboa JL, Martínez-Jaramillo S, van der Leij M, Thurner S. The multi-layer network nature of systemic risk and its implications for the costs of financial crises. Journal of Financial Stability. 2015;20:70–81.

[13] Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. Journal of Complex Networks. 2014;2:203–271.

[14] Boccaletti S, Bianconi G, Criado R, et al. The structure and dynamics of multilayer networks. Physics Reports. 2014;544:1–122.

[15] Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. Journal of Statistical Mechanics: Theory and Experiment. 2008;2008:P10008.

[16] Edmonds J. Optimum branchings. Journal of Research of the National Bureau of Standards B. 1967;71B:233–240.

[17] Veličković P, Cucurull G, Casanova A, Romero A, Liò P, Bengio Y. Graph attention networks. International Conference on Learning Representations. 2018.

[18] Brody S, Alon U, Yahav E. How attentive are graph attention networks? International Conference on Learning Representations. 2022.