ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P23

Complex-Network Modeling and Theoretical Algorithms for Secure Cross-Domain Big-Data Flows

Chang'an Ma

Ningbo University of Finance & Economics, Ningbo, Zhejiang, China.

Chang'an Ma.

machangan@nbufe.edu.cn

Abstract. Cross-domain flow can change what data means in terms of security, even when the content itself stays the same. This paper puts forward an auditable complex-network framework that brings together contextual sensitivity, typed propagation, spectral criticality, and authorization feedback into one coupled system. A temporal heterogeneous lineage graph is used to represent subjects, data objects, tasks, interfaces, and flow events. A constrained state model is built to capture the accumulation, decay, and event-driven jumps that occur over time; an explicit nonnegative matrix, rather than some opaque embedding, serves as the carrier for transmission. Under local quasi-stationarity conditions, the spectral radius of a next-generation operator gives a cascade boundary, and Perron sensitivity is applied to rank candidate channels. A finite-trust Bayesian update together with a discrete spectrum-constrained authorization problem closes the one-period-lag feedback loop. Identifiability and stability conditions are stated, a sparse monitoring algorithm is given, and layered falsification tests along with degradation rules are defined. The outcome is a mechanism-oriented theory that keeps local threshold claims separate from predictive association and industry generalization.

Keywords: cross-domain data flow; temporal heterogeneous graph; contextual sensitivity; spectral criticality; finite-trust authorization

1. Introduction

Data can be copied, derived, composed, and moved across both technical and organizational boundaries. Classification rules are able to identify intrinsic sensitivity [1, 2], but the same field may turn out to be more revealing once it gets joined with ownership changes, logistics traces, contract performance, or financing verification. The bytes of that field do not change at all, yet relational exposure, inference potential, and purpose deviation together alter the security consequence that access brings about. Secure cross-domain flow therefore calls for a dynamic account of meaning on top of a record of movement.

Existing research strands only solve parts of this problem. Provenance work represents generation, use, and derivation [3, 4], and provenance graphs have been applied to support attack detection and attribution [5]. Temporal and heterogeneous graph networks learn from typed, time-ordered interactions [6, 7], but the embeddings or attention scores they produce are not necessarily transmission parameters. Diffusion theory supplies threshold concepts [8-11], and RBAC, ABAC, usage control, and zero-trust approaches constrain authorization [12-16]. What remains missing is an auditable interface where context changes propagation, propagation yields a critical boundary, that boundary restricts authorization, and authorization in turn changes the next network state.

The theory here is organized around four maps: s to B, (B,D) to (K,R), (R,χ) to C_safe, and P to B(t+1). In this setup, s stands for context-sensitive state; B and D are transmission and recovery matrices; K is the next-generation operator; R is its spectral radius; χ is edge sensitivity; C_safe is the safe decision set; and P is a finite authorization policy. The contribution is not a loose pipeline where graph learning, a spectral score, and access control sit side by side. Each map has observable inputs, mathematical constraints, falsification tests, and a degradation rule attached to it. The threshold result is explicitly local: it is not presented as a universal phase transition or a causal effect in situations where temporal order, intervention, and confounding remain uncontrolled.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P23

Table 1. Core quantities, evidence sources, and theoretical constraints

Map	Operational role	Core constraint	Primary falsification
s → B	Context modulates typed transmission	Monotone state; nonnegative supported edges	State permutation and parameter recovery
(B,D) → (K,R)	Recovery-normalized cascade margin	Local stationarity; positive recovery	Nonlinear persistence and blocking direction
(R,χ) → C_safe	Criticality restricts feasible policies	Recompute after material masks	Freeze spectral inputs; test slack
P → B(t+1)	Authorization changes next-period support	Finite actions; one-period lag	Fix the mask; test loop stability

2. Context-Aware Temporal Network Model

2.1 Typed lineage and Contextual State

At event time t, let G_t=(V_t,E_t,T_V,T_E,τ,ψ) be a temporal heterogeneous lineage graph. The node set contains subjects, data objects, business tasks, and interfaces. A directed edge (i,j,r) records access, derivation, transfer, invocation, or business dependence; τ is there to preserve order and delay, and ψ stores purpose, privilege, task consistency, elapsed time, and other explicit evidence. Direction follows operational and provenance semantics rather than similarity. Cross-domain change is a typed event that may alter jurisdiction, purpose, interface, security domain, or contractual responsibility. Stable pseudonyms need to preserve node and edge types, relative order, reachability, and temporal intervals; if they fail to do so, the record can calibrate semantic variables but cannot identify propagation.

For object i, define z_i(t)=[a_i,e_i(t),c_i(t),p_i(t)], which represents intrinsic attributes, relational exposure, combinatorial inference potential, and purpose deviation. Its context-sensitive state is

$$s_i(t)=\sigma\left(\alpha_0+\sum_m \alpha_m z_{im}(t)+\sum_{m<n}\alpha_{mn}z_{im}(t)z_{in}(t)\right),\quad s_i(t)\in[0,1] \tag{1}$$

This state is a conditional amplification potential. It is not a legal classification level, and it is not a direct leakage probability either. Main effects are constrained to be nonnegative, while sparse interactions get accepted only when the derivative of the linear predictor with respect to each component stays nonnegative on the observed domain. Between material events, time-varying components decay toward baselines; access migration, derivation, association, or purpose change creates an explicit jump. Identification requires temporal coverage of the active variables and events, stable rank of the constrained design, and repeated contexts or controlled comparisons. If these conditions fail, interactions are removed or replaced by rule intervals rather than point estimates.

2.2 Explicit Propagation and Local Criticality

A graph encoder may propose hidden states and candidate supports, but transmission is carried only by an explicit nonnegative matrix. For relation r and admissible edge i to j, let

$$s_i(t+1)=\Pi_{[0,1]}\left\{(1-\kappa_i)s_i(t)+\kappa_i f\left[z_i(t)\right]+\xi_i(t)+\sum_q \zeta_{iq}1(t=\tau_q)\right\} \tag{2}$$

The indicator enforces lineage direction, φ contains declared edge evidence, g_r is a nondecreasing contextual modulation, and κ_r controls delay. Unsupported edges are set to zero. Relation matrices can be summed on a common state space or kept as blocks in a supra-adjacency operator. Let q_i(t) be node-risk intensity, D(t)=diag(d_i(t)) with 0<d_i≤1 be recovery, and η(t) be exogenous disturbance. In an interior event window,

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P23

$$q(t+1) \& = \Pi_{[0,1]}\Big(\big(I - D(t) + B(t)\big)q(t) + \eta(t)\Big), K(t) \& = D(t)^{-1}B(t), R(t) \& = \rho\big[K(t)\big]. \quad (3)$$

Suppose B is nonnegative, D is positive diagonal, and parameters are fixed within a quasi-stationary window, with projection being inactive. Then the unforced linearization is Schur stable if and only if R<1. This result follows because D-B is a nonsingular M-matrix exactly when the nonnegative matrix D^(-1)B has spectral radius below one [9, 10]. So R=1 serves as a local cascade boundary: R>1 indicates local amplification, not inevitable global failure. When saturation happens frequently, the Jacobian drifts, or nonlinear simulation contradicts the linear prediction, R is kept only as a warning statistic.

For a simple dominant eigenvalue, positive left and right Perron vectors u and v give the first-order channel sensitivity

$$\chi_{ijr}(t) = \frac{\partial R(t)}{\partial b_{ijr}(t)} = \frac{u_i(t)v_j(t)}{\big[\mathbf{u}(t)\ \ \mathbf{v}(t)\big]d_i(t)} \quad (4)$$

Normalized χ times the active edge weight ranks small perturbations. It is a screening measure, not a causal effect: every material deletion or rewiring must be followed by recomputation of R. The report identifies the critical strongly connected block, inbound triggers, relation composition, and whether a given intervention breaks circulation or just reduces weight. Near-repeated dominant roots invalidate individual-edge ranking; in that case relation groups or cut sets are tested directly. These qualifications keep spectral interpretation tied to directed network structure.

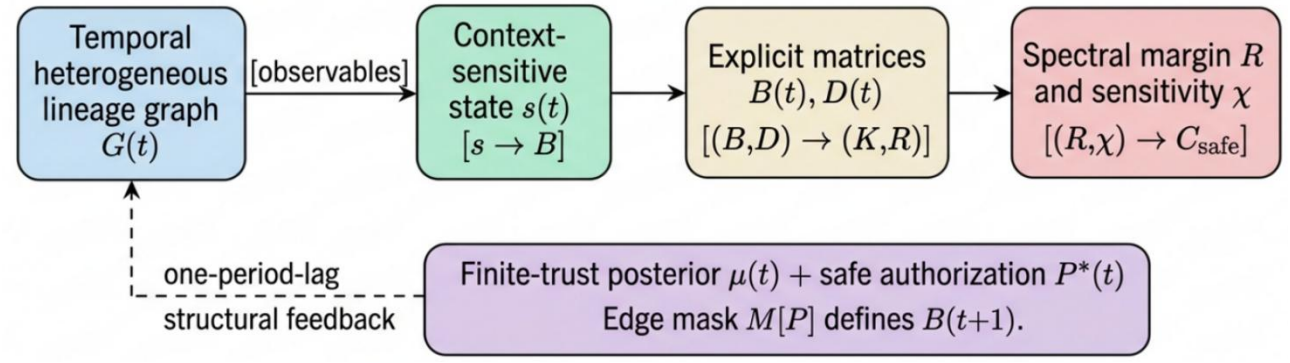


Fig. 1 Coupled interfaces for context-sensitive propagation, spectral criticality, and finite authorization feedback.

3. Theoretical Algorithms and Authorization Feedback

3.1 Sparse Estimation and Monitoring

Estimation proceeds in identifiable layers. State parameters are fitted under monotonicity and sparsity constraints; the graph encoder proposes admissible support; B is fitted under nonnegativity, direction, decay, and relation constraints; and D is calibrated from observed response or dissipation. Joint end-to-end refinement is allowed only after each layer passes parameter-recovery tests. Nested chronological validation selects interaction sparsity, relation pooling, decay, and representation dimension without future leakage. A model that has high predictive AUC but unstable propagation coefficients gets rejected for mechanistic use.

For each event batch, the Context-Spectral Flow Guard updates s, constructs admissible typed edges, evaluates B and D, decomposes the graph into strongly connected components, and computes the leading Perron pair of relevant blocks by sparse power or Arnoldi iteration. It then ranks policy-feasible masks, recomputes R after each material change, and returns the least disruptive mask that satisfies the spectral constraint. With |E| active edges and k iterations, a full sparse Perron update costs O(k|E|) time and O(|V|+|E|) memory, not counting encoder cost. Warm starts and component-local updates reduce average latency without changing the worst case.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P23

Every update records the window, support, parameter version, projection rate, critical block, eigen-residual, candidate interventions, recomputed radii, and feasibility outcome. If the information matrix is rank deficient, interactions are dropped; if relation transmission cannot be separated, channels are pooled; if the Perron root is nearly defective, block perturbations replace edge derivatives; and if local stationarity fails, nonlinear continuation and Monte Carlo stress tests become the primary evidence. The monitor therefore makes a distinction among a theorem-qualified estimate, a conservative warning bound, and a degraded output.

Authorization search is performed over a finite catalogue that comes from business roles, purpose bindings, and permitted privilege levels. The solver first eliminates decisions that violate hard policy, then applies sensitivity as a lower-cost ordering heuristic, and finally recomputes the spectral radius for every retained material mask. Branch-and-bound works well when decisions share monotone denial structure; a verified lookup table is better when the catalogue is small and latency requirements are strict. Because minimum-denial selection under a spectral constraint can be combinatorial, the system reports candidate-set size, worst-case time, and 95th-percentile time. If the solver fails to finish within the decision interval, the best previously verified safe mask is invoked, not an unverified partial optimum.

3.2 Finite-Trust Posterior and Safe Authorization

Let θ belong to a finite set of evidence-generating mechanisms, such as compliant, opportunistic, or adversarial response patterns. These are latent behavioral models, not legal labels. Given observed task completion, purpose deviation, response to intervention, and incident disposition, the posterior is

$$\mu_i(t+1,\theta)=\frac{l\left[o_i(t)\middle|\theta,a,P(t)\right]\mu_i(t,\theta)}{\sum_{\theta'} l\left[o_i(t)\middle|\theta',a,P(t)\right]\mu_i(t,\theta')} \tag{5}$$

Identification requires stable rank or detectable separation of the likelihood columns. High posterior entropy and bootstrap rank instability trigger type merging or a set-valued posterior. Strategic concealment is restricted to a finite, auditable action set such as delay, route migration, or low-intensity deviation; an arbitrary continuous disguise space would not be able to support the claimed identification or optimization.

Let C_0(t) contain policies that satisfy protection-level matching, purpose-task consistency, minimum necessity, and separation of duties. A finite policy P induces an edge mask M[P] on the predicted next-period matrix B_tilde(t+1). The safe set and selected decision are

$$C_{safe}(t)=\{P\in C_0(t):\rho[D(t)^{(-1)}\{M[P]\odot B_{tilde}(t+1)\}]\le 1+\varepsilon_t\}, \tag{6a}$$

$$P^*(t)=\arg\min_{P\in C_{\text{safe}}(t)}\left\{L_s(P;t)+\lambda_u L_u(P;t)+\lambda_v L_v(P;t)\right\} \tag{6b}$$

The dimensionless losses measure security exposure, justified-task denial, and policy switching. Weights are selected on an offline Pareto frontier and checked across time windows; mandatory denials remain hard constraints. The one-period-lag feedback B(t+1)=M[P*(t)] elementwise B_tilde(t+1) avoids simultaneous circularity. If C_safe turns out to be empty, the solver reports spectral slack and infeasibility rather than hiding the conflict by arbitrary reweighting. Finite actions and bounded loss guarantee a stage optimum, but they do not guarantee uniqueness or dynamic convergence. Around an interior fixed point, the lagged closed loop is locally stable only when the spectral radius of its full update Jacobian is below one; delayed evidence, migration, and policy switching all need to be included in this check.

4. Validation and Applicability Boundaries

Validation is layered because no single dataset can support every claim. Authorized pseudonymized enterprise records test semantic variables, contextual jumps, and within-workflow

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P23
blocking only after a topology-compatibility audit has been carried out. Public provenance graphs test typed structure, missingness, mimicry, and attribution robustness [5]; temporal transaction graphs test scale, imbalance, and cross-time transfer [17]. Controlled mechanism simulations test parameter recovery, threshold location, intervention direction, infeasible safe sets, and closed-loop stability. Public evidence does not establish commodity-industry generalization, and records that lose type, order, reachability, or delay under anonymization are excluded from propagation estimation.

Experiments target the four interfaces. Fixing or time-permuting s tests s to B; replacing contextual modulation by a constant tests state-dependent transmission; substituting a pure graph classifier tests what value explicit B adds; freezing R and χ tests spectral feedback; and fixing M tests policy-to-structure feedback. Node deletion, edge blocking, relation permutation, temporal shuffling, and path-preserving mimicry are used to separate semantic, topological, and temporal explanations. Baselines include static sensitivity scoring, temporal and heterogeneous graph models, homogeneous diffusion, RBAC, ABAC, usage control, stateful least privilege, and a no-feedback controller.

Evaluation is chronological and scenario-aware. State evidence uses recovery error, monotonicity violations, calibration, and interaction stability; propagation evidence uses threshold error, spectral margin, peak risk, cascade depth, and duration; control evidence uses exposure, business denial, policy switches, constraint violations, solution time, and Pareto stability. H1 is rejected when context jumps cannot be recovered with sign stability. H2 is rejected when R fails to align with nonlinear persistence, critical support, or blocking direction. H3 is rejected when likelihood separation fails, safe decisions are systematically infeasible, or the closed-loop Jacobian together with simulation indicates oscillation. Predictive accuracy on its own cannot establish a mechanism claim.

Causal language is restricted to controlled or credibly identified interventions. Enterprise before-and-after comparisons require temporal precedence, an explicit counterfactual, and confounding controls; otherwise they are reported as associations. Reproducibility requires versioned schemas, typed event definitions, split rules, random seeds, constraints, stopping criteria, eigen-residuals, code for synthetic data, and a claim-to-evidence table. Sensitive enterprise data need not be released, but preprocessing logic and topology diagnostics must be independently auditable. The framework is applicable where typed lineage, response evidence, and finite authorization actions exist; without these, it degrades to semantic calibration or descriptive monitoring.

In a supply-chain financing workflow, for example, a warehouse-receipt object can acquire higher contextual state when ownership changes, logistics routes, and credit verification get associated outside the approved task. Equation (2) assigns transmission only to lineage-consistent access, derivation, transfer, and interface dependencies. The critical block may then turn out to identify a repeated verification-transfer loop rather than a single high-degree node. Candidate controls can restrict an off-purpose query, lower an interface privilege, or interrupt a relation group; each candidate control is checked for business necessity and its recomputed spectral effect. This example shows the intended use of the theory while not making an unsupported claim that one parameterization transfers unchanged across firms or sectors.

5. Summary

This paper presents a compact theory for secure cross-domain big-data flow. Context-sensitive state drives an explicit nonnegative propagation matrix; recovery-normalized spectral criticality defines a local cascade boundary; Perron sensitivity screens structural interventions; and a finite-trust, spectrum-constrained policy closes an auditable lagged loop. The framework contributes most when its assumptions are visible: direction and timing are preserved, parameters are locally identifiable, the dominant spectral structure is stable, and policy actions are finite. When those conditions fail, predefined degradation rules reduce the claim rather than forcing a precise answer. Future work should put emphasis on controlled interventions, cross-workflow transfer, and robust bounds for graphs that switch rapidly.

ISSN: 3007-9640 DOI: https://doi.org/10.70114/acmsr.2026.9.1.P23

Acknowledgements

This work was supported by the 2026 Ningbo New Think Tank Project (Research on Security Resilience Governance of Cross-Border Supply Chains of Ningbo High-Tech Manufacturing Industry Driven by Digital Technology).

References

[1] Standing Committee of the National People's Congress. Data Security Law of the People's Republic of China. Beijing, China, 2021.

[2] Standardization Administration of China. GB/T 43697-2024: Data security technology—Rules for data classification and grading. Beijing, China, 2024.

[3] Buneman P, Khanna S, Tan W C. Why and where: A characterization of data provenance. In: Van den Bussche J, Vianu V, eds. Database Theory—ICDT 2001. Berlin: Springer, 2001, 316-330.

[4] Lebo T, Sahoo S, McGuinness D, eds. PROV-O: The PROV Ontology. W3C Recommendation, 2013.

[5] Han X, Pasquier T, Bates A, et al. Unicorn: Runtime provenance-based detector for advanced persistent threats. In: NDSS Symposium. San Diego, CA, 2020.

[6] Rossi E, Chamberlain B, Frasca F, et al. Temporal graph networks for deep learning on dynamic graphs. arXiv:2006.10637, 2020.

[7] Hu Z, Dong Y, Wang K, Sun Y. Heterogeneous graph transformer. In: Proceedings of The Web Conference 2020. New York: ACM, 2020, 2704-2710.

[8] Newman M E J. Networks. 2nd ed. Oxford: Oxford University Press, 2018.

[9] Berman A, Plemmons R J. Nonnegative Matrices in the Mathematical Sciences. Philadelphia: SIAM, 1994.

[10] Farina L, Rinaldi S. Positive Linear Systems: Theory and Applications. New York: Wiley, 2000.

[11] Luo X, Mao B, Tang Y, et al. On the stability and dynamics of positive systems on time-varying networks. IEEE Transactions on Automatic Control, 2022, 67(9): 4866-4873.

[12] Sandhu R S, Coyne E J, Feinstein H L, Youman C E. Role-based access control models. Computer, 1996, 29(2): 38-47.

[13] Hu V C, Ferraiolo D, Kuhn R, et al. Guide to Attribute Based Access Control Definition and Considerations. NIST SP 800-162. Gaithersburg, MD: NIST, 2014.

[14] Park J, Sandhu R. The UCONABC usage control model. ACM Transactions on Information and System Security, 2004, 7(1): 128-174.

[15] Rose S, Borchert O, Mitchell S, Connelly S. Zero Trust Architecture. NIST SP 800-207. Gaithersburg, MD: NIST, 2020.

[16] Li N, Tripunitara M V. Security analysis in role-based access control. ACM Transactions on Information and System Security, 2006, 9(4): 391-420.

[17] Weber M, Domeniconi G, Chen J, et al. Anti-money laundering in Bitcoin: Experimenting with graph convolutional networks for financial forensics. arXiv:1908.02591, 2019.