Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

Data Lineage for Cyber Incident Triage in Financial Systems: Graph-Based Impact Analysis and Recovery Prioritization

Chukwudera Obumneke Anunagba (0009-0003-3949-5290)[1*], David Excel Ozowara (0009-0003-2386-2949)[2]

[1] *Barclays, Whippany, New Jersey, USA.*

[2] *Goldman Sachs, Dallas, Texas, USA.*

***Corresponding author:** Chukwudera Obumneke Anunagba, chiokemuo@gmail.com*

***Author emails:** Chukwudera Obumneke Anunagba: chiokemuo@gmail.com; David Excel Ozowara: daveozzy6@gmail.com*

***Disclaimer:** The views expressed in this article are those of the authors and do not represent the views of Barclays or Goldman Sachs.*

ARTICLEINFO

Received: 01 Sept 2025

Revised: 15 Oct 2025

Accepted: 26 Oct 2025

ABSTRACT

Cyber incidents in financial systems frequently propagate across interconnected applications, databases, payment infrastructures, identity services, and business processes, complicating the identification of root causes, affected assets, and appropriate recovery actions. Conventional incident triage approaches often prioritize isolated alerts or technical severity without sufficiently considering data provenance, dependency propagation, transaction integrity, and operational recovery constraints. This paper examines the integration of data lineage, heterogeneous knowledge graphs, causal dependency analysis, and multi-criteria decision models to support incident triage and recovery prioritization in financial environments. Building on existing research in provenance-based attack reconstruction, graph analytics, cyber resilience, and infrastructure restoration, the study develops a conceptual lineage-to-recovery decision architecture linking observed security events to affected financial services and candidate response actions. The proposed framework distinguishes observed causal relationships from inferred dependencies, incorporates evidence confidence into impact analysis, and formulates recovery prioritization in terms of business criticality, transaction integrity, service continuity, resource constraints, and residual compromise risk. An illustrative financial incident scenario and a proposed evaluation methodology demonstrate how the framework could be operationalized and assessed. The analysis identifies research gaps concerning financial transaction semantics, provenance completeness, causal fidelity, and recovery validation. The paper contributes a structured foundation for explainable, evidence-driven incident response and recovery decision support in interconnected financial systems.

Keywords: Data Lineage; Cyber Incident Triage; Financial Systems; Provenance Graphs; Graph-Based Impact Analysis; Cyber Resilience; Recovery Prioritization; Decision Support Systems.

1 INTRODUCTION: FINANCIAL-SYSTEM INCIDENT TRIAGE AS A LINEAGE AND DEPENDENCY PROBLEM

Financial-system incident triage concerns more than detecting anomalous activity. It requires determining which event initiated a disturbance, how the disturbance propagated through data and service dependencies, which business functions were exposed, and which response or recovery action will reduce loss under constrained resources. **Data lineage** supplies evidence about the origin, movement, transformation, and use of information, while **dependency analysis** connects that evidence to processes, infrastructure, services, and organizational responsibilities. Provenance tracing has been proposed for revealing the root cause, path, and results of prolonged attacks, although storage and runtime overhead constrain deployment at scale (Tan et al., 2018).

A graph-based approach links heterogeneous observations that are otherwise separated across audit logs, cloud platforms, vulnerability records, threat intelligence, transaction systems, and incident reports. Provenance graphs represent causal relations among processes, files, and other system entities; knowledge graphs add semantic

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

relations among assets, threats, vulnerabilities, and business functions (Fang et al., 2021; Noel et al., 2017). For financial institutions, the resulting investigative object is therefore a time-dependent dependency graph rather than an isolated alert. The graph must support backward reconstruction of attack origin, forward estimation of technical and business consequences, prioritization of investigative work, and selection of recovery actions. The following sections develop this lineage-to-recovery perspective.

1.1 Research Problem, Objectives, and Contributions

Despite advances in cyber threat detection, provenance analysis, and operational resilience, incident triage in financial systems remains fragmented across security monitoring, transaction processing, infrastructure management, and business continuity functions. Security monitoring platforms primarily identify suspicious activities, while financial applications maintain transaction histories and operational systems separately record service dependencies and restoration states. Consequently, incident responders may identify a compromised asset without adequately determining its downstream effects on transaction integrity, payment processing, settlement operations, or regulatory obligations. This fragmentation creates a decision gap between detecting an incident and selecting the most appropriate containment and recovery actions.

The central research problem addressed in this paper is how heterogeneous data lineage and dependency relationships can be integrated into an explainable, graph-based decision framework that supports cyber incident reconstruction, business impact estimation, and recovery prioritization under uncertainty and limited operational resources.

The study pursues four principal objectives:

1. To characterize the provenance, causal, semantic, and operational dependency relationships required to reconstruct cyber incidents affecting financial systems.

2. To develop a conceptual graph-based framework connecting security events, financial data objects, infrastructure components, business services, and recovery states.

3. To formulate interpretable impact assessment and recovery prioritization models incorporating business criticality, evidence confidence, service interdependence, and resource constraints.

4. To establish an evaluation framework for assessing causal reconstruction quality, triage effectiveness, operational scalability, and recovery decision quality in financial environments.

The paper contributes a unified lineage-to-recovery architecture that integrates previously separate analytical perspectives: security provenance, heterogeneous dependency modeling, financial business impact assessment, and resource-constrained restoration planning. Its contribution is primarily conceptual and methodological. Rather than claiming a new experimentally validated detection algorithm, the study specifies how established graph-based methods can be combined, governed, and evaluated to support recovery-aware incident triage in interconnected financial systems.

2 FROM DATA PROVENANCE TO SECURITY-RELEVANT LINEAGE

2.1 Provenance as Evidence of Data Origin, Transformation, and Responsibility

Data provenance describes information about how a digital result was produced, including its source, transformations, and processing context. The provenance literature distinguishes uses such as accountability, reproducibility, and process debugging, while also showing that provenance forms and collection methods vary with runtime, scalability, and privacy requirements (Herschel et al., 2017). In incident triage, these dimensions acquire a security interpretation: origin identifies the initial source or entry point, transformation records processing that may have altered content or state, and responsibility associates an operation with a process, account, host, service, or team. A transaction record, for example, can be connected to the application process that created it, the data store that accepted it, and the administrative or machine identity that authorized the operation, provided the relevant events were captured and retained.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

This evidence changes the investigative question from whether a record appears anomalous to how its current state was produced. A provenance-aware representation can connect a sensitive object to preceding reads, writes, executions, transfers, and privilege changes. The connection is valuable for distinguishing primary compromise from downstream contamination, and for identifying records that require validation, rollback, or quarantine. Provenance-aware knowledge representation research emphasizes that ordinary semantic triples do not automatically preserve the origin or authority of the statements they express; provenance must be attached to assertions and retained at an appropriate granularity (Sikos & Philp, 2020). That requirement is especially relevant when incident decisions depend on conflicting observations from independent sensors.

Cloud forensics makes the evidence problem more difficult because data may originate in multiple services, move across hosts, and change during transit or processing. A survey of cloud provenance identifies source tracing, usage records, updates, and processing history as means of locating malicious vulnerabilities, while emphasizing the difficulty of capturing volatile evidence before it is overwritten (Abiodun et al., 2022). Provenance therefore has two simultaneous functions. It supports causal reconstruction after an event, and it constrains the confidence that can be assigned to an inference when collection was incomplete, delayed, or altered. A graph edge without collection time, source identity, or integrity context should be treated as weaker evidence than an edge supported by an authenticated audit event.

2.2 Cloud, Banking, and Online-Service Contexts for Lineage-Aware Investigation

Financial systems combine heterogeneous environments: core banking applications, payment and settlement services, identity platforms, customer channels, cloud-hosted components, and monitoring infrastructure. Their dependencies are also organizational. A service may depend on a database, an identity provider, an external interface, and a team responsible for remediation. Banking incident analysis using a knowledge graph demonstrates the value of integrating incident knowledge over time and extracting entities and relations from unstructured banking incident material; in that study, an improved entity-recognition model achieved a 2%-9% improvement in F1 over comparison models on 1,000 incident materials (Kang & Cheung, 2023). The result supports semantic incident organization, but it does not by itself establish causal lineage for individual transactions or demonstrate recovery outcomes.

Online-service research supplies a related operational setting. Incident triage is often revised as engineers exchange information, rather than decided once at intake. Across eight real online service systems, reassignment rates ranged from 5.43% to 68.26%, and discussions before correct assignment reached an average of 11.32 items (J. Chen et al., 2019). These findings indicate that lineage must be updated as evidence changes responsibility and scope. A graph that records only the initial alert cannot represent the evolving interpretation of an incident. Graph-based incident aggregation addresses a complementary problem by encoding temporal and topological correlations among cascading incidents and using fine-grained monitoring indicators to estimate the complete failure scope (Z. Chen et al., 2021).

The banking context adds consequences that are not captured by infrastructure topology alone. A compromised authentication service may affect account access, while a corrupted ledger process may affect balances, reporting, settlement, and regulatory obligations through different dependency paths. Business-oriented threat modeling argues that adversaries can target supporting platforms along a dependency sequence, allowing disruption to transmit from one business function to another even when conventional attack descriptions focus on lateral movement through infrastructure (Ao et al., 2025). Lineage-aware investigation must therefore connect technical events to business objects and functions without treating semantic association as proof of causation. The distinction between observed event, inferred dependency, and predicted consequence should remain explicit in the graph.

2.3 Operational Requirements for Trustworthy and Actionable Provenance

A usable provenance system must preserve temporal order, causal direction, source identity, event context, and uncertainty. It must also limit evidence growth through filtering, summarization, or query-specific materialization. Attack provenance surveys identify non-trivial space and runtime overhead as barriers to broad deployment (Tan et al., 2018). A study comparing provenance-graph database systems reports realistic generation rates of at least 1.2 MB/s and finds a trade-off between resource use in relational systems and expressive graph querying in graph

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

systems (Hänsel, 2024). These constraints make collection architecture part of triage design rather than a later implementation detail.

Security provenance also requires integrity and access control. If analysts cannot determine whether an event was collected before or after an attacker obtained administrative access, the event may support only a weak inference. If sensitive transaction or customer data is copied into an unrestricted graph, the investigation itself creates a confidentiality risk. The provenance survey identifies security and privacy as persistent challenges, alongside evidence complexity and volatile data (Abiodun et al., 2022). Practical designs should therefore separate sensitive payloads from metadata where possible, record evidence provenance for graph assertions, and expose only the minimum information needed for a decision.

The final requirement is actionability. Provenance should answer bounded questions such as which processes could have modified a record, which assets depend on a compromised identity service, or which recovery action would restore the greatest protected service. AIQL demonstrates this orientation by combining a domain-specific model and query language with execution optimizations for attack investigation over large monitoring streams; its evaluation used 857 GB and 2.5 billion events, with faster execution than several general-purpose alternatives (Gao et al., 2018). The implication is not that one query language solves financial triage, but that lineage must be organized around investigative questions. These requirements motivate causal and heterogeneous graph construction.

3 CONSTRUCTING CAUSAL AND HETEROGENEOUS GRAPHS FOR INCIDENT RECONSTRUCTION

3.1 System Audit Events, Dependency Graphs, and Attack Provenance

System provenance graphs represent entities such as processes and files as nodes and their dependencies as directed edges. Given a point-of-interest event, causal analysis can trace preceding dependencies to expose an attack sequence, but naïve construction produces graphs too large for direct analyst inspection (Fang et al., 2021). The graph is therefore an intermediate representation, not automatically an explanation. Its usefulness depends on whether edges preserve the causal relations that distinguish attack activity from routine system behavior.

Audit events provide the basic observations. Process creation, file access, network communication, credential use, and data modification can be represented as typed relations with timestamps and source metadata. The graph can then encode both direct operations and derived dependencies, such as a process depending on a file it read or a service depending on an identity provider. Provenance-based intrusion detection research treats these relations as the basis for detecting attacks and reconstructing scenarios (Suárez & Tapiador, 2025). A financial implementation would need to add transaction objects, account or customer domains, application components, control points, and business services, while preserving the distinction between observed operations and inferred dependencies.

Causality is not equivalent to temporal adjacency. An event that follows another may be unrelated, while a delayed effect may depend on an earlier action through a stored object or credential. Attack-alert correlation addresses this problem by matching the consequences of earlier alerts with the prerequisites of later attacks, thereby constructing scenarios from logical relations rather than alert order alone (Ning et al., 2002). Attack graphs similarly represent exploit dependencies through directed structures, including AND/OR conditions and cycles that require explicit handling (Zenitani, 2022). These models are complementary: alert correlation joins abstract attack steps, while provenance graphs record system-level evidence for the steps.

The graph should preserve the distinctions among causal, temporal, semantic, and administrative edges. A causal edge may indicate that one operation enabled another; a temporal edge may indicate only ordering; a semantic edge may link an alert to a known technique; and an ownership edge may identify the responsible team. Conflating them causes analysts to treat a hypothesis as an observed fact. Provenance-aware representation addresses this problem by attaching contextual information to statements and by supporting different levels of granularity (Sikos & Philp, 2020). In practice, each edge should carry its relation type, observation source, time interval, confidence, and any transformation used to derive it.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

APT detection illustrates why contextual graph structure matters. APT-KGL constructs a heterogeneous provenance graph containing system entities and contextual information, learns representations offline, and performs online detection on small local graph samples; it also creates virtual training samples from open threat knowledge to address limited labeled data (T. Chen et al., 2022). A separate provenance-graph method extracts execution traces and applies online adaptive metric learning, reporting an average 11.3% increase in detection accuracy and an average 18.3% increase in true-positive rate over its baselines (Akbar et al., 2022). These results concern simulated or benchmarked attack settings rather than banking transactions, but they support the proposition that relational context can improve detection beyond isolated event features.

3.2 Knowledge Graphs and Multi-Relational Representations of Cyber Incidents

A **knowledge graph** adds concepts and relations that are not reducible to system-event causality. Nodes may represent assets, vulnerabilities, threat actors, tactics, controls, incidents, business services, and recovery states. Edges can express "affects," "exploits," "depends on," "owned by," "supports," or "observed in," with provenance attached to each assertion. Cybersecurity knowledge-graph research describes a progression from ontologies and semantic triples toward richer graphs that support reasoning across heterogeneous security information (Zhao et al., 2023). Threat-intelligence frameworks likewise use formal ontologies to combine threat descriptions, network configurations, observables, and vulnerability information for relevance and risk analysis (Qamar et al., 2017).

The advantage of this representation is cross-domain alignment. An event from an endpoint sensor can be linked to a vulnerability, an attack technique, a cloud service, a payment function, and a responsible team. CyGraph illustrates this architecture through a multi-relational property graph that combines network infrastructure, security posture, cyber threats, and mission dependencies, enabling queries for risky multi-step paths from threats to vulnerable hosts and important cyber assets (Noel et al., 2017). For financial systems, the same pattern can connect an alert to transaction processing, customer authentication, settlement, fraud controls, and recovery alternatives. Such links permit impact analysis that respects both technical reachability and service purpose.

The representation must accommodate relations involving more than two entities. Threat-intelligence reports may describe an actor using a technique against an asset through a vulnerability under particular conditions. Hypergraph research addresses this limitation by representing multiple relations among security entities and reports improved entity and relation extraction on 685 threat-intelligence reports (Jia et al., 2024). This is useful where a simple edge would lose context, although extraction accuracy does not guarantee that the resulting relation is causally valid in a live financial environment. Entity extraction from banking incident text similarly supports knowledge integration, but it remains distinct from event-level reconstruction (Kang & Cheung, 2023).

Graph learning can identify patterns that rules miss, but its outputs require interpretation. APT-KGL and other provenance methods show that graph representations can support local classification and anomaly detection (Mei et al., 2021; T. Chen et al., 2022). MGDA combines self-supervised graph representation learning with contextual analysis and rule-based pattern matching, producing scenario graphs with higher-level attack semantics; its evaluation across three datasets reported average precision of 97.58% and an F1-score of 98.03% (Cui et al., 2025). These metrics establish performance on the evaluated datasets, not general reliability in financial operations. A triage architecture should retain the supporting subgraph, source events, and alternative explanations rather than expose only a classification label.

3.3 Formal Representation of the Financial Incident Dependency Graph

To operationalize the integration of data provenance and business dependencies, the proposed framework represents the financial environment as a temporal, heterogeneous, directed property graph:

$$G_t = (V_t, E_t, \mathcal{R}, \Phi_t)$$

where G_t denotes the graph state at time (t), V_t is the set of entities, E_t is the set of directed relationships, $\mathcal{R}$ is the set of permissible relationship types, and Φ_t contains node- and edge-level attributes, including timestamps, evidence sources, confidence values, operational states, and integrity indicators.

The node set is partitioned into six principal categories:

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

$$V_t = V_A \cup V_D \cup V_S \cup V_B \cup V_I \cup V_R$$

where V_A represents infrastructure and application assets, V_D represents financial data objects and transactions, V_S represents technical services, V_B represents business functions, V_I represents security incidents and evidence, and V_R represents recovery resources and operational states.

A graph relationship is represented as a typed, time-bounded assertion:

$$e = (u, v, r, t_s, t_e, c, p)$$

where (u) and (v) identify the source and destination entities, $r \in \mathcal{R}$ specifies the relationship type, t_s and t_e define the applicable temporal interval, $c \in [0,1]$ represents confidence in the assertion, and (p) identifies the underlying evidence provenance.

The relation vocabulary includes observed operations such as read, write, execute, authenticate, and transmit, together with derived relationships such as depends on, supports, affected by, and recoverable through. Observed and inferred relationships must remain distinguishable because semantic association or temporal proximity does not independently establish causation.

Graph updates are performed incrementally as evidence becomes available. When an incident occurs, backward traversal identifies possible initiating events and enabling conditions, whereas forward traversal estimates affected data objects, services, and business processes. Traversal is constrained by temporal consistency, relation semantics, and evidence confidence to reduce spurious propagation.

This formal representation provides a common foundation for incident reconstruction, impact assessment, prioritization, and recovery optimization while retaining the information necessary to explain how a recommended action was derived.

The resulting graph can be understood as a layered structure:

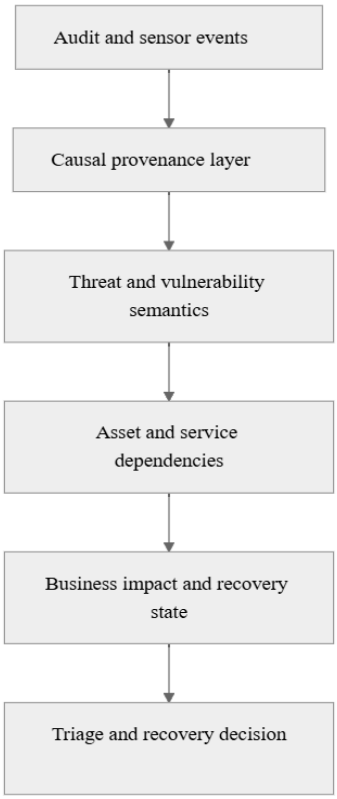


Figure 1. Layered Data Lineage and Graph-Based Cyber Incident Triage Architecture for Financial Systems.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

Figure 1 illustrates the conceptual integration of security event collection, causal provenance reconstruction, threat and vulnerability interpretation, asset and service dependency mapping, business impact assessment, and recovery decision support. The architecture connects raw observational evidence to financial business consequences through a series of analytical layers. The representation emphasizes that recovery decisions should be derived from traceable evidence and contextual dependency relationships rather than isolated alert severity. In a practical implementation, the layers would exchange updates bidirectionally as new evidence modifies incident scope, confidence, and recovery readiness.

3.4 Causal Fidelity, Graph Abstraction, and Scalable Investigation

Causal fidelity is the extent to which a reconstructed graph preserves the attack relations that matter for investigation. Node-level precision and recall may indicate whether entities were detected, but they can miss altered or missing causal structure. SDTED was proposed to compare provenance graphs while preserving structure and depth, and its evaluation on DARPA Engagement attack scenarios found that it detected changes in causal relationships where node-level metrics were less effective (Suárez & Tapiador, 2025). This distinction matters for financial incidents because a graph containing the correct hosts but an incorrect path may lead investigators toward the wrong account, credential, or recovery action.

Abstraction is therefore necessary, but it must be controlled. DEPIMPACT assigns discriminative weights to dependency edges, propagates impact backward from a point of interest, and then performs forward analysis from highly ranked entry nodes. On 150 million real system-auditing events and DARPA data, it reduced graphs of approximately 1,000,000 edges to approximately 234 edges while preserving attack sequences, a 4,611-fold reduction (Fang et al., 2021). Such reduction can make analyst review feasible, but an omitted edge may represent the only connection to a contaminated transaction or privileged account. Summaries should retain expansion paths and state the criteria used to remove or down-weight evidence.

Query systems address a related scalability problem. AIQL uses attack-specific query primitives and execution strategies to investigate large monitoring datasets, while ProGQL (Shao et al., 2025) introduces constrained traversal, edge weighting, value propagation, and incremental graph search without requiring full in-memory materialization. These approaches suggest that scalable investigation depends on semantics-aware storage and search rather than simply selecting a faster general-purpose database. The choice of abstraction should follow the question: entry-point discovery needs backward traversal, contamination analysis needs forward traversal, and responsibility assignment needs ownership and service relations.

Evaluation must test more than detection accuracy. TDLens compared five provenance-graph detection approaches using approximately 40 GB of host-level logs and six simulated attack scenarios, with the purpose of identifying which attack types methods detect and why (Mei et al., 2022). MGDA evaluated detection and scenario reconstruction across real-world and simulated network attacks (Cui et al., 2025). These studies offer evidence about algorithmic behavior, but they do not establish preservation of financial transaction semantics, regulatory evidence requirements, or analyst trust. A robust evaluation should therefore measure causal reconstruction, time to triage, graph reduction, false exclusions, provenance completeness, and the quality of resulting impact and recovery decisions.

4 GRAPH-BASED IMPACT ANALYSIS ACROSS TECHNICAL, BUSINESS, AND CYBER-PHYSICAL DEPENDENCIES

4.1 Backward Propagation from Alerts, Points of Interest, and Compromised Assets

Impact analysis begins from an anchor: an alert, suspicious transaction, modified file, compromised identity, unavailable service, or other point of interest. Backward propagation asks which prior entities and operations could have enabled the observation. This direction is useful for identifying entry points, credential theft, initial access, and the earliest trustworthy evidence. DEPIMPACT formalizes the approach by assigning edge weights, propagating dependency impact backward from the point of interest, and using forward analysis from high-ranked entry nodes to remove unrelated dependencies (Fang et al., 2021). The method demonstrates that directionality can reduce investigation volume while retaining attack structure.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

Backward traversal should be constrained by relation type and time. A process that wrote a file may be relevant to a later execution, whereas a process that merely coexisted on the same host may not be. Attack-alert correlation provides a complementary constraint by matching prerequisites and consequences among alerts (Ning et al., 2002). Attack graphs add exploit conditions and alternative paths, allowing analysts to ask whether a vulnerability and network position jointly enabled access rather than treating either feature as sufficient (Zenitani, 2022). These relations can produce a ranked set of hypotheses instead of a single asserted root cause.

The same method applies to financial objects. If an anomalous payment is the point of interest, backward analysis may examine the initiating session, authentication event, authorization decision, application process, data read, and preceding administrative changes. If several transactions share a credential, device, or service dependency, the graph can identify a broader exposure set. The inference remains conditional on data completeness: missing endpoint telemetry, unlogged service-to-service calls, or overwritten cloud evidence can interrupt the path. Provenance surveys identify this evidence-capture problem as a central limitation, especially for volatile data (Abiodun et al., 2022).

Backward propagation also supports responsibility assignment. A graph can distinguish the team operating a compromised service from the team responsible for the vulnerable component or the account owner whose credentials were abused. Continuous triage research shows that initial assignments are often revised through discussion and new evidence (J. Chen et al., 2019). An evidence graph can make these revisions explicit by adding new observations and changing confidence in candidate causes rather than overwriting the original assessment. This preserves an audit trail for why the incident scope and owner changed.

4.2 Forward Propagation of Data, Service, and Infrastructure Consequences

Forward propagation estimates what may have been affected after a compromised event. Starting from an asset or operation, the graph follows data flows, service dependencies, privilege relations, and infrastructure connections. The result can identify exposed records, dependent applications, degraded controls, and services whose operation may be unsafe. Provenance-based damage assessment for cloud intrusions emphasizes context and multi-stage attacks, treating the extent of damage to a victim virtual machine as an input to response (Arshad et al., 2013). The approach supports the principle that response should depend on affected context rather than on an intrusion label alone.

Data propagation and service propagation are related but distinct. A compromised process may alter records that flow into reporting or settlement, while a failed authentication service may prevent access without corrupting data. The graph should therefore model read, write, transmit, execute, authenticate, and depend-on relations separately. A business dependency model describes threat transmission from one supporting business to another when a prior function becomes abnormal (Ao et al., 2025). This model extends technical reachability by asking which business processes cannot operate normally when an upstream function is impaired.

Cyber-physical systems require another layer because cyber events can change physical states or flows. A hierarchical flow model combines functional hierarchy, cyber-physical interaction within a station, dependencies among stations, and physical network topology to assess impact spreading in a gas supply system (Q. Zhu et al., 2019). Dependency-based risk assessment similarly identifies attack paths to critical physical components by modeling dependencies between cyber and physical elements (Akbarzadeh & Katsikas, 2022). For financial infrastructure, the physical layer may include data centers, network facilities, power systems, and communications links rather than industrial actuators, but the analytical requirement is similar: service consequences may emerge through dependencies that are absent from the initial cyber alert.

Forward propagation should represent uncertainty and alternative states. An edge may indicate a confirmed data transfer, a possible dependency, or a predicted consequence based on a model. If all edges are treated equally, the graph exaggerates certainty and can trigger unnecessary isolation or restoration. Graph analytics for cyber-physical resilience uses multilayer knowledge-graph models to identify components that may generate critical failures and compares graph metrics for locating such points (Dagnas et al., 2025). The evidence supports graph-based critical-point analysis, while also showing why metric choice affects which components are judged important.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

4.3 Business Dependencies and the Financial-System Impact Surface

The **financial-system impact surface** is the set of technical assets, data objects, services, business functions, stakeholders, and obligations that may be affected by an incident. It cannot be inferred from host criticality alone. A customer-facing channel may be replaceable if a redundant channel exists, whereas an apparently ordinary data service may support multiple settlement or compliance processes. Business dependency modeling explicitly represents platforms and business functions together and characterizes threat effects as they transmit through dependencies (Ao et al., 2025). This provides a conceptual basis for associating technical compromise with operational consequences.

The graph should distinguish direct impact from dependency-mediated impact. Direct impact includes a modified ledger record, unavailable database, or compromised administrator account. Mediated impact includes delayed settlement, incorrect reporting, unavailable fraud screening, or loss of customer access caused by the direct event. The distinction supports different response actions: data validation may address integrity risk, traffic restriction may address propagation, and alternate processing may preserve service availability. Incident knowledge graphs can accumulate prior incident entities and relations, but their extracted semantics should be linked to current observations before they influence a live decision (Kang & Cheung, 2023).

Business criticality also depends on time and concentration. A service supporting many functions may have high structural centrality, but a less central service may be the only path for a time-sensitive obligation. Infrastructure vulnerability screening uses interconnected digraphs and multiattribute utility theory to rank scenarios according to stakeholder values and the consequences of losing services (Apostolakis & Lemon, 2005). This supports a multiattribute impact surface rather than a single technical score. The attributes may include affected customers, transaction integrity, service duration, substitutability, regulatory exposure, and recovery effort, provided each is defined and evidenced.

4.4 Multi-Dimensional Financial Incident Impact Assessment

Financial incident impact should be evaluated using multiple dimensions rather than infrastructure connectivity or technical severity alone. A graph-based impact assessment model can combine the criticality of affected services, the integrity of financial records, the extent of transaction exposure, operational availability, and regulatory consequences.

For an incident (i), a proposed normalized impact score is defined as:

$$I_i = w_B B_i + w_D D_i + w_T T_i + w_A A_i + w_R R_i$$

subject to:

$$\sum_{k \in \{B,D,T,A,R\}} w_k = 1, \qquad w_k \geq 0$$

where:

- B_i represents business-service criticality.
- D_i represents financial data integrity impact.
- T_i represents transaction and customer exposure.
- A_i represents operational availability impact.
- R_i represents regulatory and compliance consequences.
- w_k represents an institution-defined importance weight.

All component scores are normalized to the interval [0, 1], with larger values representing greater impact. The weights should be determined through documented institutional risk criteria, stakeholder consultation, or a calibrated multi-criteria decision process.

The graph-derived exposure of an affected business function (b) may be represented as:

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

$$X_b = \max_{p \in \mathcal{P}_{ib}} \prod_{e \in p} q_e$$

where $\mathcal{P}_{ib}$ is the set of eligible dependency paths from incident (i) to business function (b), and $q_e \in [0,1]$is a calibrated conditional propagation factor associated with edge (e). This formulation is applicable only under the simplifying assumption that propagation factors along an individual path can be multiplied. When dependencies are correlated, more appropriate probabilistic graphical models or scenario-based calculations are required.

Importantly, impact severity and evidence confidence should not be conflated. A potentially catastrophic but weakly evidenced compromise should be represented as a high-consequence scenario with substantial uncertainty rather than automatically receiving a low priority. The model should therefore report estimated impact alongside evidence confidence, unresolved assumptions, and plausible upper- and lower-impact bounds.

The proposed approach enables analysts to determine not only which assets are connected to an incident but also which financial functions may experience the greatest consequential disruption. Such estimates provide an interpretable basis for escalation, evidence collection, containment, and recovery decisions.

A graph can connect these attributes to recovery choices. For example, a compromised authentication dependency may lead to customer access restrictions, while an alternate verified authentication path may reduce the projected service loss. A compromised data store may require isolation even when this reduces availability because forward propagation indicates integrity risk. The impact graph thus becomes a decision graph, in which each candidate action changes reachable services, evidence quality, and residual exposure. This transition leads to triage methods that rank incidents and investigative actions under uncertainty.

5 INCIDENT TRIAGE AND PRIORITIZATION UNDER RESOURCE AND UNCERTAINTY CONSTRAINTS

5.1 Alert Correlation, Continuous Triage, and Root-Cause Diagnosis

Incident triage allocates scarce investigative and response capacity while the incident state remains incomplete. Independent alerts create a poor decision surface because repeated symptoms can obscure a common cause. Alert-correlation methods construct attack scenarios by matching prerequisites and consequences, reducing the volume of unrelated alerts while retaining relationships among attack steps (Ning et al., 2002). Graph-based incident aggregation extends this logic to online services by encoding topological and temporal correlations among cascading incidents and using monitoring indicators to recover failure scope (Z. Chen et al., 2021).

The triage state should be updated as new evidence arrives. Continuous incident triage research shows that assignment can change repeatedly as engineers discuss an incident, with reassignment rates varying widely across online service systems (J. Chen et al., 2019). This finding argues against a static queue ordered only by the first alert. A lineage-aware queue can update severity, affected assets, candidate causes, owners, and recommended actions as graph relations are confirmed or weakened. The system should preserve previous states because changing prioritization is itself evidence about uncertainty and organizational coordination.

Root-cause diagnosis requires more than anomalous metrics. PatternMatcher identifies candidate root-cause metrics by combining coarse anomaly detection, anomaly-pattern classification, and ranking according to physical meaning and engineering needs. On four datasets containing 113 incidents from a large commercial bank, it achieved top-three average accuracy of 0.91 and was deployed in practice (Wu et al., 2021). The banking setting makes this evidence directly relevant to financial operations, although metric diagnosis does not by itself reconstruct data lineage or establish the cause of a security event. Combining metric evidence with provenance edges can distinguish an infrastructure symptom from an initiating security action.

Investigation itself is a sequential allocation problem. DISCLOSE uses probabilistic relations among adversarial tactics, techniques, and procedures, together with proximity in an attack life-cycle model, to optimize which investigative step should be taken next under constraints (Nisioti et al., 2021). The approach was evaluated through a case involving 31 TTPs, interviews with cybersecurity professionals, and ATT&CK STIX and CVSS data (Nisioti et al., 2021). In a financial setting, the next step could be acquisition of endpoint evidence, validation of transaction

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

integrity, isolation of an identity provider, or examination of a service dependency. The graph supplies candidate actions; constrained optimization determines which action offers the best information or risk reduction for available time and personnel.

5.2 Context-Sensitive Vulnerability and Incident Prioritization

Vulnerability priority cannot be separated from asset context. A vulnerability on an isolated development host may be less urgent than a lower-scored weakness on a reachable component supporting authentication or payment processing. Context-sensitive vulnerability research develops a machine-learning vulnerability priority scoring system and then uses sequential optimization to select vulnerability instances and assign them to personnel with matching skills (Hore et al., 2022). The study reports better prioritization than rule-based methods and another literature method on a real-world vulnerability dataset, but its endpoint is mitigation allocation rather than confirmed incident reduction.

A provenance and dependency graph supplies the contextual variables required for such prioritization. These include reachability from threat sources, position on paths to protected assets, privilege relationships, data sensitivity, active exploitation evidence, business criticality, redundancy, and available containment options. CyGraph demonstrates how a graph knowledge base can identify multi-step paths from threats through vulnerable hosts to important cyber assets and associate risky traffic paths with filtering devices and rules (Noel et al., 2017). The method supports both proactive remediation and reactive mitigation, which are often separated in operational workflows.

Incident priority should also incorporate evidence confidence. A high-impact consequence supported by a confirmed write to a settlement database differs from one inferred through a weak dependency relation. Conversely, low-confidence evidence may justify rapid collection if the cost of delay is high. A multi-factor prioritization methodology integrates technical severity with national criticality and additional attributes such as economic consequences, functional consequences, recovery capability, and system specifics (Uzlov et al., 2025). Its staged process includes registration, verification, correlation, criticality assessment, and tie-breaking, offering a useful structure for maintaining uncertainty rather than hiding it in one score.

Resource allocation adds personnel and timing constraints. The organization may have too few analysts, limited access to specialized systems, or a narrow window before volatile evidence disappears. The provenance survey identifies evidence complexity and volatile data capture as operational problems (Abiodun et al., 2022). A priority score should therefore combine expected harm, exploit or propagation likelihood, evidence decay, investigation cost, and action feasibility. The weights should be explicit and revisable because financial institutions may value integrity, availability, customer access, and regulatory reporting differently across incidents.

5.3 Evidence-Aware Cyber Incident Triage Priority Model

While impact assessment estimates the consequences of an incident, incident triage determines the urgency with which investigative and containment resources should be assigned. These objectives are related but distinct. A high-impact event may require immediate response because of confirmed compromise, while another event with uncertain exposure may require urgent evidence acquisition before a containment decision can be justified.

The proposed framework therefore introduces an incident triage priority model:

$$P_i = \alpha I_i + \beta L_i + \gamma U_i + \delta E_i + \epsilon F_i$$

subject to:

$$\alpha + \beta + \gamma + \delta + \epsilon = 1$$

where P_i represents the priority of incident (i), I_i denotes the normalized business impact score, L_i represents the estimated likelihood of continuing compromise or propagation, U_i represents uncertainty-sensitive investigation urgency, E_i represents the urgency associated with evidence expiration or loss, and F_i represents the time sensitivity of available interventions. All components are normalized to ([0,1]).

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

The uncertainty term represents the value of resolving important unknowns, rather than treating uncertainty itself as evidence of compromise. Similarly, intervention urgency reflects the expected consequences of delaying a feasible action rather than simply rewarding actions that are easy to perform.

Priority scores are recalculated as new observations modify the incident graph, confidence assessments, or estimated consequences. Each recalculation should generate a versioned decision record containing the input values, assigned weights, supporting evidence paths, and resulting priority classification.

Priority thresholds may be mapped to institution-specific escalation categories, such as critical, high, medium, and low. However, thresholds should be calibrated against documented operational requirements, incident response objectives, and historical decision outcomes rather than selected arbitrarily.

For high-consequence incidents, predefined escalation rules may override the numerical ranking when there is credible evidence of active financial data manipulation, continuing unauthorized transfers, or imminent loss of critical payment processing. Consequently, the proposed score functions as an explainable decision-support mechanism and does not replace established incident response authority or analyst judgment.

5.4 Technical Severity Scores Versus Topology, Criticality, and Dependency Context

Technical severity scores remain useful for normalizing vulnerability characteristics, but they do not describe the graph in which a vulnerability operates. Attack-graph analysis shows that topology, configuration, vulnerability distribution, exploit dependencies, and alternative paths determine reachable attack states (Zenitani, 2022). A score that ignores these relations may rank an easily exploitable but peripheral weakness above a modest weakness on the only path to a critical service. The correct response is not to discard technical severity, but to combine it with topology and business consequences.

Infrastructure vulnerability screening illustrates this combination through interconnected directed graphs and multiattribute utility theory (Apostolakis & Lemon, 2005). Cyber-physical risk assessment similarly ranks attack paths according to the probability and consequence of reaching critical components within a dependency model (Akbarzadeh & Katsikas, 2022). These approaches differ in domain and formalism, yet they converge on a context-sensitive principle: priority belongs to the vulnerability or incident scenario whose exploitation produces the greatest relevant loss under the stated threat assumptions. The assumptions must remain visible because rankings can change when attack budgets, recovery resources, or service substitutions change.

Topology can also affect the value of a patch or containment action. Removing a vulnerability from a bottleneck may reduce many attack paths, while patching an exposed node with abundant alternate routes may have little network-wide effect. Graph-based prioritization should therefore evaluate marginal reduction in reachable critical assets, expected impact, and residual dependency risk. This is especially relevant when a financial organization must choose among patching, segmentation, credential rotation, monitoring, and temporary service restrictions. The selected action should be explainable through the paths and business functions it changes.

The distinction between vulnerability prioritization and incident prioritization must be preserved. A vulnerability is a potential enabling condition; an incident includes evidence of activity, compromise, or consequence. A high-severity unexploited vulnerability may require remediation, while a lower-severity weakness associated with confirmed unauthorized transaction modification may require immediate investigation and containment. Multi-criteria models can represent both, but their endpoints differ. This distinction prepares the move from triage scores to recovery sequences, where action ordering changes the future state of an interdependent system.

6 RECOVERY PRIORITIZATION AND CYBER RESILIENCE IN INTERDEPENDENT SYSTEMS

6.1 Restoration Sequencing for Interdependent Infrastructure and Services

Recovery prioritization determines which components, services, data stores, and controls should be restored or replaced first. In an interdependent financial environment, restoring one service can be ineffective or unsafe if its identity, network, data, or monitoring dependencies remain compromised. Research on multiple interdependent infrastructures formulates restoration as a joint problem involving damaged components, ordering, routing, and service loss over time. The power and gas restoration model uses mixed-integer programming and randomized

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

adaptive decomposition to obtain high-quality restoration solutions for large benchmark infrastructures. Its infrastructure domain differs from banking, but its sequencing logic applies to any system in which service dependencies and restoration resources interact.

Tri-level resilience models make the dependence among protection, attack, and repair explicit. One framework selects pre-event defense, identifies a worst-case attack, and determines post-event repair, showing that the best defense depends on attack and repair budgets (Ouyang & Fang, 2017). Another model integrates facility protection with recovery decisions, including repair of disrupted facilities and capacity expansion of non-interdicted facilities (Bao et al., 2017). For financial services, analogous actions include hardening privileged access, isolating compromised components, restoring validated data, activating alternate processing, and increasing capacity in unaffected channels.

Restoration order should be based on service recovery rather than asset count. Repairing a highly damaged component may produce little benefit if an upstream dependency remains unavailable, whereas restoring a shared identity, network, or data-validation function may return several services. Recovery from localized infrastructure attacks shows that preferential repair strategies can improve network recovery compared with random repair, especially where demand is uneven across locations (Hu et al., 2016). The financial analogue is uneven service demand across customer channels, transaction types, or time-sensitive obligations. The graph should therefore associate each restoration action with expected service return, residual compromise, and possible overload.

A recovery sequence also has evidentiary conditions. A database should not be restored from a snapshot merely because it is recent if lineage indicates that the snapshot may contain attacker-modified data. Recovery actions should reference trusted provenance, validation results, and dependency state. Data dependency has long been proposed for assessing damage from malicious transactions and supporting recovery of affected information (Haraty & Zeitunlian, 2007). The graph can identify records and downstream computations requiring revalidation, while the recovery optimizer can order those actions according to business loss and available capacity.

6.2 Integrating Protection, Restoration, and Adaptive Flow Redistribution

Recovery is not limited to repairing failed components. During disruption, services may be rerouted, workloads redistributed, or capacity temporarily increased. A model integrating protection, restoration, and adaptive flow redistribution treats these decisions jointly because redistribution can overload surviving components and change network reliability (H. Zhu et al., 2021). The framework proposes a resilience metric that reflects reliability changes under redistribution and uses tri-level optimization to coordinate pre-event and post-event decisions. In a financial system, flow may represent transaction requests, authentication traffic, settlement messages, or data-processing workloads rather than electricity or material supply.

Adaptive redistribution can preserve service while increasing exposure. Routing customers to an alternate channel may maintain availability but concentrate traffic on a component with weaker monitoring or limited fraud controls. Redirecting settlement processing may preserve throughput while creating reconciliation risk. The recovery graph should therefore encode capacity, integrity requirements, control coverage, and dependency status for each alternative path. A service is not restored safely if it is reachable but cannot satisfy the controls attached to its function.

Protection decisions affect later recovery options. Building redundancy or protecting weak components can reduce the loss caused by an attack, but the preferred strategy changes with attack and repair budgets (Ouyang & Fang, 2017). Interdependent infrastructure analysis likewise finds that considering dependencies from both attacker and defender perspectives changes the resilience strategy (Ouyang, 2017). This supports a lifecycle model in which lineage and dependency graphs are maintained before, during, and after incidents. Pre-event graphs describe possible paths; incident graphs add observed evidence; recovery graphs represent remaining safe paths and restoration actions.

Adaptive redistribution also requires feedback. As components are restored, traffic can move again; as new evidence identifies compromise, an apparently available route may become unsafe. Continuous triage provides the

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

operational mechanism for revising the graph, while optimization recalculates the action sequence. The resulting process is a control loop rather than a one-time plan:

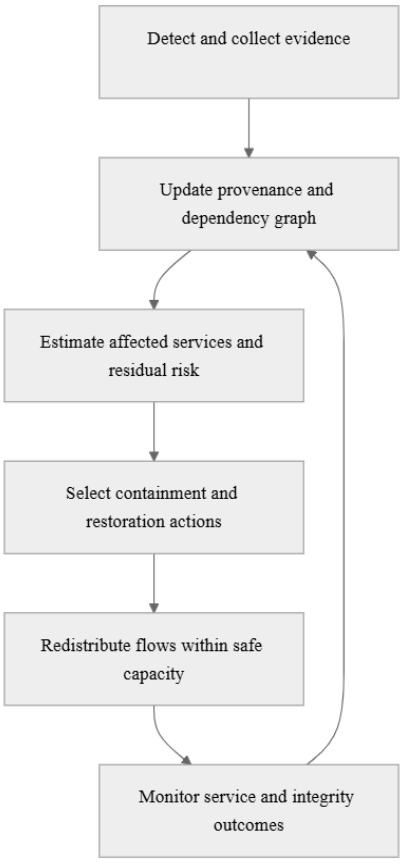


Figure 2. Closed-Loop Incident Triage and Recovery Optimization Process.

Figure 2 presents the iterative relationship between incident evidence collection, provenance graph maintenance, affected-service estimation, containment and restoration planning, workload redistribution, and recovery verification. The feedback mechanism enables the framework to revise impact estimates and recommended actions as new observations become available or the operational state changes. This closed-loop design is particularly important for financial systems, where restoring service availability without validating transaction integrity or downstream dependencies may create additional operational risk.

6.3 Resilience Optimization and Decision Support for Financial-Service Continuity

Cyber resilience concerns preparation, absorption, recovery, and adaptation after adverse cyber effects (AlHidaifi et al., 2024). A lineage-to-recovery system gives these stages a common representation. Preparation uses dependency paths to identify weak components and important services. Absorption uses provenance and monitoring to isolate affected activity while sustaining safe functions. Recovery uses validated data and dependency-aware sequencing. Adaptation uses post-incident evidence to change controls, redundancy, and investigation rules.

Optimization provides a way to compare recovery portfolios rather than selecting one action at a time. Decision-support research models uncertain threat rates, countermeasure costs, and asset impacts, then uses a genetic algorithm to search for portfolios with different cost-risk trade-offs (Rees et al., 2011). Tri-level resilience models similarly compare defense, attack, and repair decisions under changing budgets (Ouyang & Fang, 2017). These approaches show why financial continuity cannot be optimized from technical severity alone. The objective may

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

combine service loss over time, transaction integrity, customer impact, regulatory obligations, recovery cost, and residual compromise.

6.4 Dependency-Constrained Recovery Optimization Model

Recovery decisions in financial systems should minimize service disruption while ensuring that restored components satisfy integrity, security, and dependency requirements. Because individual recovery actions may influence multiple interconnected services, restoration should be represented as a constrained optimization problem rather than an isolated component-ranking exercise.

Let x(a,t) ∈ {0, 1} indicate whether recovery action (a) is initiated during time interval (t). A conceptual recovery objective is:

$$\min_{x} J(x) = \sum_{t=1}^{T} \left[\sum_{s \in S} W_s \, L_{s,t}(x) + \lambda C_t(x) + \mu Q_t(x) \right]$$

where *(J(x))* represents total recovery loss, (S) is the set of financial services, W_s represents the business importance of service (s), $L_{s,t}(x)$ represents its normalized service-loss level during interval (t), $C_t(x)$ is the incremental cost of recovery actions initiated during that interval, and $Q_t(x)$ is the estimated residual security and transaction-integrity risk. The parameters λ and μ establish the relative importance of financial expenditure and residual risk.

The recovery plan must satisfy several operational constraints.

Resource capacity constraint

$$\sum_{a \in A} r_{a,k} \, x_{a,t} \leq R_{k,t}$$

where $r_{a,k}$ represents resource type (k) required by recovery action (a), and $R_{k,t}$ represents the available resource capacity during interval (t).

Dependency readiness constraint

$$z_{s,t} \leq z_{d,t} \qquad \forall d \in D(s)$$

where $z_{s,t}$ indicates whether service (s) is operationally enabled at time (t), and $D(s)$ contains its mandatory dependencies. This constraint assumes that the dependency has no verified substitute; alternative paths must be represented explicitly.

Integrity validation constraint

$$z_{s,t} \leq v_{s,t}$$

where *(v_{s,t})* indicates successful validation of the security and data-integrity requirements applicable to service (s).

Recovery action constraint

$$\sum_{t=1}^{T} x_{a,t} \leq 1$$

for recovery actions that may be initiated at most once within the planning horizon.

Additional scheduling variables and constraints are necessary to model action duration, resource occupancy across multiple intervals, alternative recovery methods, service capacity, and precedence among restoration activities.

The optimization process must therefore consider both the restoration benefit and the conditions under which a service can be safely returned to operation. A critical payment application should not be activated merely because its infrastructure is available if the associated transaction database remains unverified or the authentication service is still compromised.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

For operational deployment, optimization could be implemented using mixed-integer linear programming, constrained heuristic search, or rolling-horizon optimization, depending on graph size, action complexity, and response-time requirements. Under uncertainty, scenario-based or robust optimization can represent alternative compromise states, restoration durations, and service dependencies.

The resulting recovery plan should provide a prioritized action sequence, the supporting dependency paths, anticipated service improvements, resource requirements, residual risks, and conditions that require plan revision. This transforms lineage analysis from retrospective forensic reconstruction into an actionable component of financial cyber-resilience planning.

The model should include uncertainty in both graph evidence and system response. A candidate restoration action may have uncertain duration, uncertain capacity, or uncertain contamination risk. Scenario analysis can represent alternative attack paths and recovery states, while sensitivity analysis can show whether the recommended action changes when criticality or resource assumptions change. The vulnerability-screening literature uses stakeholder-valued attributes to convert service consequences into decision priorities (Apostolakis & Lemon, 2005). This suggests that financial institutions should make value judgments explicit rather than burying them in unexplained weights.

A practical decision-support architecture should present ranked actions with their supporting paths. For each action, analysts need to see the initiating evidence, affected assets, dependent services, expected service restoration, residual risks, required skills, and conditions that would invalidate the recommendation. Investigation optimization can identify the next evidence-collection step (Nisioti et al., 2021), while restoration optimization can select repair and redistribution sequences (H. Zhu et al., 2021). Linking these outputs prevents a common failure in which detection, impact assessment, and recovery operate as separate queues.

The evidence base remains stronger for generic cloud, online-service, cyber-physical, and critical-infrastructure settings than for end-to-end financial transaction recovery. Financial deployment therefore requires controlled validation against transaction integrity, customer access, settlement dependencies, and audit obligations. The most defensible architecture treats graph recommendations as decision support, preserves analyst overrides, and records the evidence and assumptions behind each action. In that form, data lineage becomes a bridge from traceable incident evidence to recovery-aware prioritization rather than a separate forensic repository.

7 TOWARD A UNIFIED LINEAGE-TO-RECOVERY DECISION ARCHITECTURE

7.1 Linking Detection Evidence to Impact Scores and Recovery Actions

A unified architecture should convert **lineage evidence** into an ordered decision chain: detected event, inferred causal path, affected data or service, business impact, and recovery action. Provenance graphs expose origins, transformations, and dependencies, while alert correlation connects prerequisites and consequences into attack scenarios (Abiodun et al., 2022; Ning et al., 2002). This evidence can support impact scores that combine technical compromise, data integrity, service interruption, transaction exposure, and recovery effort rather than treating alert severity as the endpoint.

The score should remain conditional on evidence quality and system state. Backward propagation can identify likely entry points, while forward analysis estimates which dependent assets and services may be affected (Fang et al., 2021). A recovery recommendation should therefore display its supporting path, confidence, affected business function, expected restoration sequence, and residual compromise risk. Continuous triage is appropriate because incident assignments and causal interpretations change as engineers contribute new information (J. Chen et al., 2019).

7.2 Combining Provenance Graphs with Optimization and Multi-Criteria Decision Models

Provenance graphs can provide the state representation for optimization, but they do not determine institutional priorities by themselves. Multi-relational graph models already connect infrastructure, vulnerabilities, threats, and mission dependencies for resilience analysis (Noel et al., 2017). An optimization layer can then select investigation steps, mitigation actions, or restoration sequences subject to analyst capacity, system constraints,

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

cost, and time. DISCLOSE demonstrates how probabilistic relations among adversarial techniques can guide the next forensic action under investigation constraints (Nisioti et al., 2021).

Recovery models add objectives that detection systems usually omit. Interdependent-infrastructure research optimizes repair order and service loss over time, while tri-level formulations distinguish protection, attack, and post-event repair decisions (Ouyang & Fang, 2017). For financial systems, a multi-criteria model should represent transaction integrity, customer access, settlement continuity, regulatory duties, restoration duration, and contamination risk. Stakeholder-valued utility models offer a precedent for making these value choices explicit rather than embedding them in opaque weights (Apostolakis & Lemon, 2005). Scenario analysis should test whether rankings change when causal confidence, service criticality, repair capacity, or attack assumptions vary.

7.3 Governance, Explainability, and Human Oversight in Triage Decisions

Governance requires provenance retention, access control, versioned graph transformations, and an auditable record of recommendations and overrides. Provenance research identifies runtime, scalability, privacy, and representation as deployment constraints, not secondary documentation concerns (Herschel et al., 2017). Explainability should be structural: analysts should inspect the event chain, dependency edges, assumptions, and excluded alternatives behind each priority. Attack graphs can improve comprehension, although their effectiveness depends on users' technical backgrounds and the representation used (Lallie et al., 2017).

Human oversight remains necessary when evidence is incomplete or recovery actions can alter live transaction processing. Automated rankings should recommend, not authorize, disruptive isolation, rollback, or restoration. NIST-oriented incident-response guidance places detection, response, and recovery within continuing risk-management activity (Nelson et al., 2025). Institutions should record analyst rationale, uncertainty, and post-incident corrections so that the graph and decision model become reviewable operational records.

7.4 Financial Data Integrity and Regulatory Governance Requirements

Financial incident recovery introduces assurance requirements that extend beyond restoring application availability. A recovered service must preserve the correctness, completeness, authorization, and traceability of financial transactions. Consequently, the proposed framework should incorporate explicit integrity-validation checkpoints before previously compromised financial data or services are returned to production.

For transaction-processing systems, validation may include comparison of transaction identifiers, authorization records, ledger balances, database audit trails, settlement messages, and reconciliation outputs. Where relevant, integrity verification should examine whether transaction modifications were authorized, whether downstream systems consumed affected records, and whether recovery operations preserved legitimate transaction ordering and accounting consistency.

Data lineage also supports evidentiary accountability by documenting how incident conclusions were reached, which evidence sources supported those conclusions, and which transformations were applied during graph construction. Evidence records should be protected against unauthorized modification, governed by retention and access-control policies, and associated with collection and integrity-verification metadata.

The governance architecture should support applicable regulatory and contractual requirements without assuming that all financial institutions are subject to identical obligations. Requirements may differ according to jurisdiction, institutional classification, payment infrastructure, data category, and incident consequences.

Automated recovery recommendations should therefore include explicit authorization boundaries, evidence-preservation requirements, integrity-verification prerequisites, and escalation conditions. Actions involving transaction rollback, financial record modification, or restoration of critical payment services should require appropriate human approval and documented validation.

By incorporating these controls, the lineage-to-recovery framework connects cyber resilience with financial accountability and auditability, reducing the risk that technically successful restoration produces inconsistent or untrusted financial records.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

8 ILLUSTRATIVE FINANCIAL CYBER INCIDENT SCENARIO AND FRAMEWORK APPLICATION

8.1 Scenario Description

To demonstrate the practical application of the proposed lineage-to-recovery architecture, consider a hypothetical cyber incident affecting a financial institution operating an online banking platform, an authentication service, a payment processing application, a transaction database, and a settlement reporting system. The scenario is illustrative and does not represent a recorded security incident or experimentally validated deployment.

The incident begins when an unauthorized actor obtains privileged application credentials and initiates suspicious activity through the institution's payment processing environment. The actor uses the compromised identity to interact with a payment service and modify selected transaction records before the affected data is consumed by downstream settlement and reconciliation processes.

Conventional monitoring may generate several independent alerts, including abnormal privileged authentication, unusual database modification activity, suspicious transaction patterns, and reconciliation discrepancies. Without integrated lineage, these alerts may be investigated separately, delaying identification of the shared initiating condition and the scope of downstream exposure.

8.1.1 Step 1: Evidence Collection and Graph Construction

The framework collects available authentication events, application audit trails, database transaction records, network communications, payment processing logs, and settlement service metadata. These observations are represented as typed entities and relationships within the temporal financial incident graph.

The initial graph connects the compromised identity to the authenticated application session, the payment processing service, the modified database records, and the downstream business functions that consume those records. Each relationship retains its observation source, timestamp, and evidence confidence.

8.1.2 Step 2: Backward Causal Reconstruction

Backward analysis begins with the anomalous financial transaction and traverses relevant write, execution, authentication, and service relationships. The objective is to identify the earliest evidence-supported unauthorized activity and distinguish the initiating compromise from subsequent symptoms.

For example, the investigation may identify that multiple suspicious transaction modifications originated from the same application session and were associated with a privileged credential. This relationship would support further investigation of credential compromise without independently proving how the credential was acquired.

8.1.3 Step 3: Forward Business Impact Analysis

Forward traversal identifies other records, services, and business functions potentially affected by the suspicious activity. The analysis examines whether modified payment records were consumed by settlement, reconciliation, customer reporting, or downstream financial analytics processes.

The framework distinguishes confirmed record modifications from possible downstream contamination. Services that only read validated records should not automatically be classified as compromised merely because they are connected through a dependency graph.

Impact assessment considers transaction integrity, financial exposure, business-service criticality, availability consequences, and compliance obligations. The resulting impact profile identifies which functions require urgent containment, investigation, or validation.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

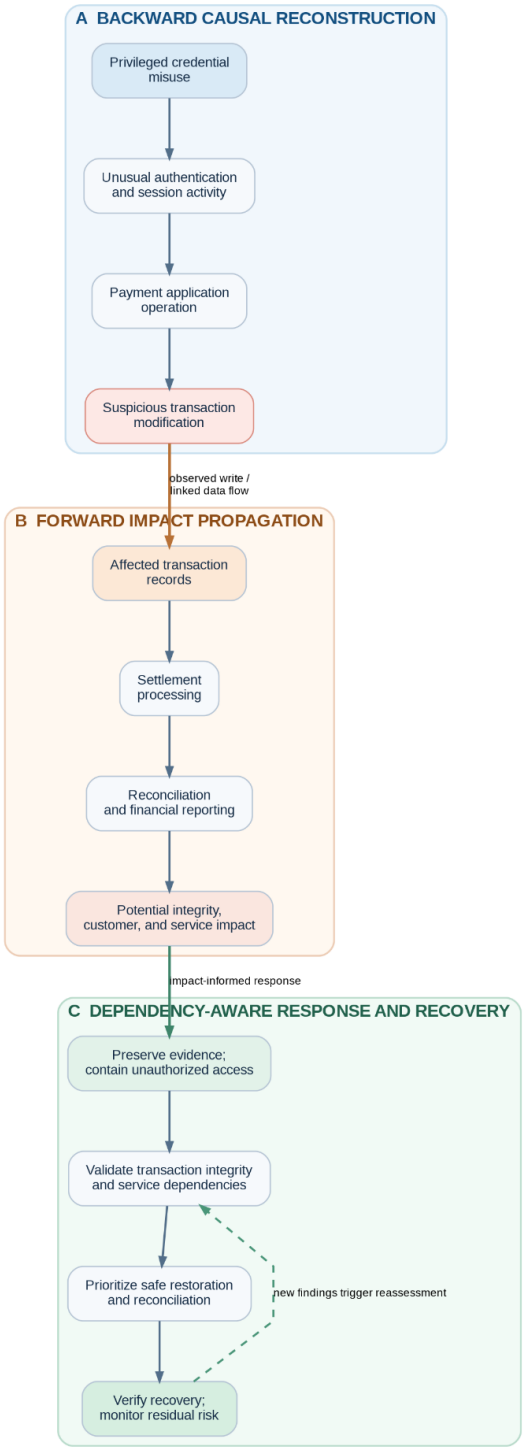


Figure 3. Financial Cyber Incident Lineage, Impact Propagation, and Dependency-Aware Recovery Framework.

The figure illustrates the integration of backward causal reconstruction, forward financial impact propagation, and dependency-aware cyber incident recovery within a unified graph-based framework. The backward reconstruction stage identifies potential initiating events through privileged credential activity, authentication records, payment application operations, and suspicious transaction modifications. The forward propagation stage traces affected financial records and their dependencies on settlement processing, reconciliation, and reporting functions to

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

estimate potential operational and financial consequences. The recovery stage connects the assessed impact to evidence preservation, containment, transaction integrity validation, prioritized service restoration, and post-recovery monitoring. The feedback mechanism enables continuous reassessment as new evidence becomes available, supporting explainable incident triage and recovery decisions.

As illustrated in Figure 3, the proposed framework integrates three interdependent analytical processes to support financial cyber incident investigation and recovery. First, backward causal reconstruction traces suspicious financial transactions through relevant application operations, authentication events, and potential credential misuse to identify evidence-supported initiating conditions. Second, forward impact propagation examines how affected transaction records may influence settlement operations, financial reconciliation, reporting systems, and other dependent business functions. This process distinguishes directly observed modifications from inferred downstream consequences, reducing the risk of overstating incident scope.

Third, dependency-aware recovery planning translates the resulting impact assessment into a sequence of containment, integrity validation, restoration, and verification activities. Recovery actions are evaluated against service dependencies and transaction integrity requirements to ensure that restored components do not reintroduce compromised information or disrupt unaffected financial processes. The framework incorporates continuous feedback, allowing new forensic observations and restoration outcomes to update the incident graph and influence subsequent decisions. This integrated approach establishes a traceable relationship between cybersecurity evidence, financial business impact, and operational recovery priorities, while preserving human oversight over consequential response actions.

8.1.4 Step 4: Incident Triage and Containment

The triage engine combines the graph-derived impact profile with active-compromise indicators, investigation urgency, evidence retention requirements, and time-sensitive response conditions.

The incident response team may prioritize the preservation of transaction logs and evidence, termination of unauthorized sessions, revocation or rotation of compromised credentials, and restriction of affected payment processing operations.

Containment actions must be evaluated against service dependencies to avoid creating unnecessary disruption in unaffected channels. Any decision to restrict transaction processing remains subject to authorized operational and security personnel.

8.1.5 Step 5: Recovery Dependency Analysis

Following containment, the recovery graph determines which components can safely resume operation. Recovery prerequisites may include confirmation that privileged access is secured, identification of affected transaction records, validation of database integrity, and verification of downstream settlement consistency.

The framework identifies valid recovery sequences based on these prerequisites. For example, restoring customer-facing payment processing before validating the transaction database could reintroduce corrupted information into downstream systems. Conversely, restoring a verified identity service and clean transaction data may enable multiple business functions to resume safely.

8.1.6 Step 6: Recovery Verification and Continuous Updating

Recovery completion is evaluated against both technical and business criteria. Technical verification includes service health, access control validation, database consistency, and monitoring coverage. Business verification includes payment accuracy, reconciliation completeness, settlement readiness, and confirmed availability of customer-facing services.

New evidence may alter the estimated compromise scope or invalidate a previously selected restoration path. The framework therefore updates the incident and recovery graphs as investigation and remediation progress.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

8.2 Scenario Implications

This hypothetical scenario illustrates the proposed framework's end-to-end decision logic: security evidence is converted into a causal hypothesis, linked to affected financial data and services, translated into business impact estimates, and used to support a dependency-aware recovery plan.

The scenario does not establish that the framework reduces incident response time or financial loss. Those claims require controlled evaluation against conventional incident response processes and verified operational outcomes.

Table 1. Graph inputs and decision outputs by investigation stage.

Investigation stage	Graph input	Decision output
Alert detection	Authentication and payment alerts	Initial incident scope
Causal reconstruction	Session and database provenance	Candidate initiating events
Impact propagation	Transaction and service dependencies	Potentially affected functions
Incident triage	Impact, evidence, urgency	Investigation and containment priority
Recovery planning	Service prerequisites and integrity state	Feasible restoration sequence
Verification	Reconciliation and system health results	Evidence-supported recovery completion

9 RESEARCH GAPS IN FINANCIAL-SYSTEM DEPLOYMENT AND EVALUATION

9.1 Limited Direct Evidence from Banking and Financial-Transaction Environments

The literature supports provenance-based detection and dependency analysis, but direct evidence from banking operations remains limited. A banking incident study applied root-cause metric identification to 113 incidents from a large commercial bank and reported a top-three average accuracy of 0.91, demonstrating practical relevance for diagnosis rather than recovery prioritization (Wu et al., 2021). A separate study modeled banking incident knowledge from 1,000 incident materials, focusing on entity extraction and knowledge integration rather than live transaction consequences (Kang & Cheung, 2023). These studies address operational diagnosis and incident semantics, but neither establishes that lineage-derived priorities improve payment integrity, settlement continuity, customer outcomes, or regulatory reporting.

Most evaluations instead use enterprise hosts, cloud services, simulated attacks, or non-financial critical infrastructure. Provenance detectors report strong results on attack datasets, yet their settings differ from institutions where transactions cross applications, legal entities, ledgers, and external networks (Akbar et al., 2022; T. Chen et al., 2022). Cloud incident aggregation has been deployed in online-service environments, but service incidents are not equivalent to compromises involving mutable financial records (Z. Chen et al., 2021). Research therefore needs controlled banking testbeds and retrospective production studies that connect technical evidence to verified business loss and recovery milestones.

9.2 Incomplete Integration of Data Lineage, Business Criticality, and Recovery State

Current methods generally emphasize one layer: provenance for causal reconstruction, knowledge graphs for semantic integration, vulnerability models for exposure, or operations research for restoration. Business-

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

dependency models show how disruption can propagate between supporting platforms and services (Ao et al., 2025), while resilience optimization shows that the preferred protection strategy changes when repair resources and recovery actions are included (Ouyang & Fang, 2017). Few studies connect these representations into a time-varying graph that records data lineage, business criticality, compromise confidence, restoration readiness, and service status together.

This omission matters because the same technical asset may have different priorities before containment, during forensic preservation, and after a clean recovery path is available. A financial institution needs models that distinguish suspected exposure from confirmed corruption, degraded service from unsafe service, and reversible processing delay from irreversible transaction error. Restoration research offers objective functions for service loss and repair ordering, but those models do not encode evidentiary uncertainty or lineage-specific integrity claims. Future work should define interoperable schemas and decision rules that preserve these distinctions.

9.3 Evaluation Challenges: Provenance Quality, Timeliness, Privacy, and Generalizability

Evaluation must measure more than alert precision or graph classification. Provenance systems face evidence loss, volatile data, storage overhead, and privacy constraints (Abiodun et al., 2022), while graph database studies report trade-offs between query expressiveness and resource consumption (Hänsel, 2024). Causal fidelity also requires metrics that compare reconstructed structure, because node-level agreement can miss altered relationships between events (Suárez & Tapiador, 2025). Banking evaluations should therefore report capture completeness, event latency, graph reduction, causal accuracy, analyst time, decision stability, and recovery outcomes.

Generalizability is another unresolved issue. Studies using simulated attacks or open threat knowledge may improve coverage, but they cannot establish performance across proprietary architectures, changing transaction workflows, or institution-specific controls (Nisioti et al., 2021; T. Chen et al., 2022). Privacy-preserving evaluation must permit reproducibility without exposing customer data, credentials, or commercially sensitive topology. Comparative benchmarks should include clean, degraded, and compromised recovery states, with independent validation of business impact and analyst decisions.

9.4 Proposed Experimental Evaluation and Validation Framework

The effectiveness of the lineage-to-recovery architecture should be evaluated through controlled experiments incorporating financial-system dependencies, representative attack scenarios, and measurable recovery outcomes. Since the present study proposes a conceptual framework, the following methodology defines a validation agenda rather than reporting completed experimental results.

Evaluation Environment

A representative financial-system testbed should contain an identity service, payment processing application, transaction database, reconciliation component, monitoring infrastructure, and business continuity controls. Synthetic or appropriately de-identified transaction records should be used to avoid exposing confidential financial information.

The evaluation should include scenarios involving unauthorized transaction modification, privileged credential compromise, payment service disruption, transaction-data contamination, and multi-service dependency failures. Each scenario should be accompanied by a documented reference timeline identifying known initiating events, affected records, relevant dependency paths, and expected restoration conditions.

Comparative Baselines

The proposed framework should be evaluated against three alternative approaches:

1. **Alert-severity baseline:** Incidents are prioritized using alert severity and conventional asset-level attributes without explicit lineage traversal.

2. **Topology-only baseline:** Impact is estimated using service and infrastructure dependencies without transaction-level provenance or evidence-confidence modeling.

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Journal of Information Systems Engineering and Management
2025, 10(61s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ Research Article

3. **Provenance-only baseline:** Incidents are reconstructed using recorded causal relationships without incorporating business criticality or recovery optimization.

A fourth comparison may use manual analyst-directed triage, subject to consistent scenario access, analyst experience, and resource constraints.

Evaluation Metrics

Performance should be measured across four dimensions.

Incident reconstruction quality: Precision, recall, and F1-score for identifying relevant causal relationships, affected entities, and confirmed attack paths, supplemented by a graph-structure fidelity measure.

Operational effectiveness: Time to identify the initiating event, time to establish the affected service scope, number of investigative actions, analyst effort, and graph query latency.

Business impact accuracy: Precision and recall of affected financial services, false-exclusion rate for contaminated transaction records, and deviation between predicted and verified service consequences.

Recovery performance: Time to verified service restoration, cumulative business-service disruption, resource utilization, residual compromise indicators, and proportion of restored services satisfying integrity-validation requirements.

Experimental Procedure

Each evaluation scenario should be executed under predefined conditions, and the resulting audit events and dependency data should be supplied to the comparison methods. Where feasible, analysts should be blinded to scenario ground truth and method presentation should be randomized to reduce learning and ordering effects.

Experiments should vary graph size, evidence completeness, event delay, compromised asset criticality, and available recovery resources. Particular attention should be given to incomplete evidence because financial incident investigations frequently involve missing audit events or uncertain external-service dependencies.

For each scenario, the framework should produce a causal reconstruction, an affected-service estimate, a triage ranking, and a proposed recovery sequence. These outputs should be compared with the documented scenario ground truth and objectively measured recovery states.

Statistical and Sensitivity Analysis

Results should be reported with descriptive statistics and appropriate uncertainty intervals. When repeated measurements or paired scenarios are available, statistical comparisons should account for the experimental design and dependence among observations.

Ablation studies should separately remove temporal constraints, provenance confidence, business criticality, and recovery dependency modeling to determine their contributions to overall performance. Sensitivity analysis should evaluate changes in prioritization and recovery sequences under different institutional weight assignments and missing-evidence levels.

The framework should be considered empirically supported only where observed results demonstrate reliable improvements against the stated baselines. In particular, detection accuracy alone would not establish improved financial recovery outcomes, and successful service restart would not establish recovery unless required security and transaction-integrity conditions are verified.

10 CONCLUSION: FROM TRACEABLE INCIDENT EVIDENCE TO RECOVERY-AWARE PRIORITIZATION

Effective cyber incident triage in financial systems requires more than the identification of suspicious activities or the assignment of technical severity scores. Financial applications operate through interconnected dependencies involving transaction databases, authentication services, payment platforms, settlement processes, infrastructure

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

components, and regulatory obligations. Consequently, a security incident originating in one technical component may produce operational consequences across multiple financial services and data-processing stages.

This paper has examined how data provenance, heterogeneous dependency graphs, causal incident reconstruction, business impact assessment, and recovery optimization can be integrated into a unified lineage-to-recovery decision architecture. The proposed conceptual framework connects observed security events with affected financial data objects and business services, enabling the systematic examination of incident origin, propagation pathways, operational exposure, and candidate recovery actions.

The study contributes a formal temporal graph representation, interpretable impact and triage scoring formulations, and a dependency-constrained recovery optimization model. Together, these elements establish a methodological foundation for translating heterogeneous incident evidence into explainable investigation and recovery priorities. An illustrative financial incident scenario demonstrates how the framework could support the progression from alert detection to causal investigation, downstream impact assessment, containment, restoration planning, and service validation.

The analysis also identifies important limitations. Provenance collection may be incomplete, graph relationships may contain uncertain or inferred dependencies, and recovery optimization may depend on assumptions about component integrity, restoration duration, and institutional priorities. Moreover, existing empirical evidence remains insufficient to establish the operational superiority of an integrated lineage-to-recovery approach across diverse financial environments.

Future research should implement the proposed architecture within representative financial transaction environments and evaluate it against severity-based, topology-only, and provenance-only baselines. Emphasis should be placed on causal reconstruction accuracy, transaction contamination detection, investigation efficiency, recovery sequencing, graph scalability, and verified business-service restoration.

Ultimately, the value of data lineage in cyber incident response extends beyond retrospective forensic analysis. When integrated with business dependency modeling, financial integrity controls, and governed recovery decision support, lineage can provide a foundation for more transparent, context-sensitive, and recovery-aware cybersecurity operations.

DECLARATIONS

Author contributions: Chukwudera Obumneke Anunagba: Conceptualization, Writing – original draft, Writing – review & editing. David Excel Ozowara: Conceptualization, Writing – original draft, Writing – review & editing.

Funding: This research received no external funding.

Competing interests: The authors declare no competing interests.

Data availability: No new data were generated or analysed in this study.

Use of generative AI: Generative AI tools were used to assist with drafting and language editing. The authors reviewed and edited all content and take full responsibility for the manuscript.

REFERENCES

[1] Abiodun, O. I., Alawida, M., Omolara, A. E., & Alabdulatif, A. (2022). Data provenance for cloud forensic investigations, security, challenges, solutions and future perspectives: A survey. *Journal of King Saud University - Computer and Information Sciences*, *34*(10), 10217–10245. https://doi.org/10.1016/j.jksuci.2022.10.018

[2] Akbar, K. A., Wang, Y., Ayoade, G., Gao, Y., Singhal, A., Khan, L. R., Thuraisingham, B. M., & Jee, K. (2022). Advanced persistent threat detection using data provenance and metric learning. *IEEE Transactions on Dependable and Secure Computing*, *20*(5), 3957–3969. https://doi.org/10.1109/tdsc.2022.3221789

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[3] Akbarzadeh, A., & Katsikas, S. (2022). Dependency-based security risk assessment for cyber-physical systems. *International Journal of Information Security*, *22*(3), 563–578. https://doi.org/10.1007/s10207-022-00608-4

[4] AlHidaifi, S. M., Asghar, M. R., & Ansari, I. S. (2024). A survey on cyber resilience: Key strategies, research challenges, and future directions. *ACM Computing Surveys*, *56*(8), 1–48. https://doi.org/10.1145/3649218

[5] Ao, S., Fang, B., Xiao, X., & Zhang, H. (2025). Threat ripple model: A model to characterize business-oriented attacks based on business dependencies. *Security and Safety*. *https://doi.org/10.1051/sands/2025003*

[6] Apostolakis, G., & Lemon, D. M. (2005). A screening methodology for the identification and ranking of infrastructure vulnerabilities due to terrorism. *Risk Analysis*, *25*(2), 361–376. https://doi.org/10.1111/j.1539-6924.2005.00595.x

[7] Arshad, J., Azad, M. A., Jokhio, I. A., & Townend, P. (2013). Intrusion damage assessment for multi-stage attacks for clouds. *IET Communications*, *7*(12), 1304–1315. https://doi.org/10.1049/iet-com.2012.0707

[8] Bao, S., Zhang, C., Ouyang, M., & Miao, L. (2017). An integrated tri-level model for enhancing the resilience of facilities against intentional attacks. *Annals of Operations Research*, *283*(1–2), 87–117. https://doi.org/10.1007/s10479-017-2705-y

[9] Chen, J., He, X., Lin, Q., Zhang, H., Hao, D., Gao, F., Xu, Z., Dang, Y., & Zhang, D. (2019). Continuous incident triage for large-scale online service systems. In *IEEE/ACM International Conference on Automated Software Engineering (ASE)* (pp. 364–375). https://doi.org/10.1109/ase.2019.00042

[10] Chen, T., Dong, C., Lv, M., Song, Q., Liu, H., Zhu, T., Xu, K., Chen, L., Ji, S., & Fan, Y. (2022). APT-KGL: An intelligent APT detection system based on threat knowledge and heterogeneous provenance graph learning. *IEEE Transactions on Dependable and Secure Computing*, 1–15. https://doi.org/10.1109/tdsc.2022.3229472

[11] Chen, Z., Liu, J., Su, Y., Zhang, H., Wen, X., Ling, X., Yang, Y., & Lyu, M. R.-T. (2021). Graph-based incident aggregation for large-scale online service systems. In *IEEE/ACM International Conference on Automated Software Engineering (ASE)* (pp. 430–442). https://doi.org/10.1109/ase51524.2021.9678746

[12] Cui, M., Jiang, Z., Li, S. C., Ma, C., Zhang, K., Yang, P., & Feng, H. (2025). MGDA: A provenance graph-based framework for threat detection and attack scenario reconstruction. *Computer Networks*, *274*, 111806. https://doi.org/10.1016/j.comnet.2025.111806

[13] Dagnas, R., Barbeau, M., García-Alfaro, J., & Yaich, R. (2025). Graph analytics for cyber-physical system resilience quantification [Preprint]. In *ArXiv.org*. https://doi.org/10.48550/arxiv.2504.02120

[14] Fang, P., Gao, P., Ayday, E., Jee, K., Wang, T., Ye, Y., Liu, Z., & Xiao, X. (2021). Back-Propagating system dependency impact for attack investigation. In *Zenodo (CERN European Organization for Nuclear Research)*. https://doi.org/10.5281/zenodo.5559213

[15] Gao, P., Xiao, X., Li, Z., Jee, K., Xu, F., Kulkarni, S. R., & Mittal, P. (2018). AIQL: Enabling efficient attack investigation from system monitoring data [Preprint]. In *arXiv (Cornell University)* (pp. 113–125). https://doi.org/10.48550/arxiv.1806.02290

[16] Hänsel, S.-O. (2024). Untersuchung von datenbanksystemen zur verwaltung von provenance graphen im kontext der IT-Sicherheit. In *SerWisS (Hochschule Hannover)*. https://doi.org/10.25968/opus-3232

[17] Haraty, R., & Zeitunlian, A. (2007). *Damage Assessment and Recovery from Malicious Transactions Using Data Dependency for Defensive Information Warfare*.

[18] Herschel, M., Diestelkämper, R., & Ben Lahmar, H. (2017). A survey on provenance: What for? what form? what from? *The VLDB Journal*, *26*(6), 881–906. https://doi.org/10.1007/s00778-017-0486-1

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[19] Hore, S., Moomtaheen, F., Shah, A., & Ou, X. (2022). Towards optimal triage and mitigation of context-sensitive cyber vulnerabilities. *IEEE Transactions on Dependable and Secure Computing*, *20*(2), 1270–1285. https://doi.org/10.1109/tdsc.2022.3152164

[20] Hu, F., Yeung, C. H., Yang, S., Wang, W., & Zeng, A. (2016). Recovery of infrastructure networks after localised attacks. *Scientific Reports*, *6*(1), 24522. https://doi.org/10.1038/srep24522

[21] Jia, J., Li, Y., Wang, Y., & Sang, A. (2024). Hyper attack graph: Constructing a hypergraph for cyber threat intelligence analysis. *Computers & Security*, *149*, 104194. https://doi.org/10.1016/j.cose.2024.104194

[22] Kang, W., & Cheung, C. F. (2023). Banking system incidents analysis using knowledge graph. *International Journal of Knowledge and Systems Science*, *14*(1), 1–23. https://doi.org/10.4018/ijkss.325794

[23] Lallie, H. S., Debattista, K., & Bal, J. (2017). An empirical evaluation of the effectiveness of attack graphs and fault trees in cyber-attack perception. *IEEE Transactions on Information Forensics and Security*, *13*(5), 1110–1122. https://doi.org/10.1109/tifs.2017.2771238

[24] Mei, R., Yan, H., Han, Z., & Jiang, J. (2021). CTSCOPY: Hunting cyber threats within enterprise via provenance graph-based analysis. In *IEEE International Conference on Software Quality, Reliability and Security (QRS)* (pp. 28–39). https://doi.org/10.1109/qrs54544.2021.00014

[25] Mei, R., Yan, H., Wang, Q., Han, Z., & Lyu, Z. (2022). TDLens: Toward an empirical evaluation of provenance graph-based approach to cyber threat detection. *China Communications*, *19*(10), 102–115. https://doi.org/10.23919/jcc.2022.00.028

[26] Nelson, A., Rekhi, S., Souppaya, M., & Scarfone, K. (2025). *Incident response recommendations and considerations for cybersecurity risk management :* National Institute of Standards and Technology (U.S.). https://doi.org/10.6028/nist.sp.800-61r3

[27] Ning, P., Cui, Y. F., & Reeves, D. S. (2002). Constructing attack scenarios through correlation of intrusion alerts. In *Proceedings of the ACM Conference on Computer and Communications Security* (pp. 245–254). https://doi.org/10.1145/586110.586144

[28] Nisioti, A., Loukas, G., Lászka, Á., & Panaousis, E. (2021). Data-Driven decision support for optimizing cyber forensic investigations. *IEEE Transactions on Information Forensics and Security*, *16*, 2397–2412. https://doi.org/10.1109/tifs.2021.3054966

[29] Noel, S., Bodeau, D., & McQuaid, R. (2017). *Big-Data Graph Knowledge Bases for Cyber Resilience*.

[30] Ouyang, M. (2017). A mathematical framework to optimize resilience of interdependent critical infrastructure systems under spatially localized attacks. *European Journal of Operational Research*, *262*(3), 1072–1084. https://doi.org/10.1016/j.ejor.2017.04.022

[31] Ouyang, M., & Fang, Y. (2017). A mathematical framework to optimize critical infrastructure resilience against intentional attacks. *Computer-Aided Civil and Infrastructure Engineering*, *32*(11), 909–929. https://doi.org/10.1111/mice.12252

[32] Qamar, S., Anwar, Z., Rahman, M. A., Al-Shaer, E. S., & Chu, B.-T. B. (2017). Data-driven analytics for cyber-threat intelligence and information sharing. *Computers & Security*, *67*, 35–58. https://doi.org/10.1016/j.cose.2017.02.005

[33] Rees, L. P., Deane, J. K., Rakes, T. R., & Baker, W. H. (2011). Decision support for cybersecurity risk planning. *Decision Support Systems*, *51*(3), 493–505. https://doi.org/10.1016/j.dss.2011.02.013

[34] Shao, F., Zou, J., Cao, Z., & Xiao, X. (2025). ProGQL: A provenance graph query system for cyber attack investigation. arXiv. https://doi.org/10.48550/arXiv.2510.22400

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

[35] Sikos, L. F., & Philp, D. (2020). Provenance-Aware knowledge representation: A survey of data models and contextualized knowledge graphs. *Data Science and Engineering*, *5*(3), 293–316. https://doi.org/10.1007/s41019-020-00118-0

[36] Suárez, M. S., & Tapiador, J. (2025). Attack structure matters: Causality-preserving metrics for provenance-based intrusion detection systems. *Computers & Security*, *157*, 104578. https://doi.org/10.1016/j.cose.2025.104578

[37] Tan, C. Y., Wang, Q., Wang, L., & Zhao, L. (2018). Attack provenance tracing in cyberspace: Solutions, challenges and future directions. *IEEE Network*, *33*(2), 174–180. https://doi.org/10.1109/mnet.2018.1700469

[38] Uzlov, D., Yakovlev, S., Tolstoluzka, O. G., Kopytsia, O., & Burchenko, S. (2025). Integrating CVSS, national criticality levels, and MCDA for multi-factor cyber incident prioritization. *RADIOELECTRONIC AND COMPUTER SYSTEMS*, *2025*(4), 220–235. https://doi.org/10.32620/reks.2025.4.15

[39] Wu, C., Zhao, N., Wang, L., Yang, X., Li, S., Zhang, M., Jin, X., Wen, X., Nie, X., Zhang, W., Sui, K., & Pei, D. (2021). Identifying root-cause metrics for incident diagnosis in online service systems. In *IEEE International Symposium on Software Reliability Engineering (ISSRE)* (pp. 91–102). https://doi.org/10.1109/issre52982.2021.00022

[40] Zenitani, K. (2022). Attack graph analysis: An explanatory guide. *Computers & Security*, *126*, 103081. https://doi.org/10.1016/j.cose.2022.103081

[41] Zhao, X., Jiang, R., Han, Y., Li, A., & Peng, Z. (2023). A survey on cybersecurity knowledge graph construction. *Computers & Security*, *136*, 103524. https://doi.org/10.1016/j.cose.2023.103524

[42] Zhu, H., Zhang, C., Ramírez-Márquez, J. E., Wu, S., & Monroy, R. (2021). The integration of protection, restoration, and adaptive flow redistribution in building resilient networked critical infrastructures against intentional attacks. *IEEE Systems Journal*, *15*(2), 2959–2970. https://doi.org/10.1109/jsyst.2020.3039466

[43] Zhu, Q., Qin, Y., Zhou, C., & Fei, L. (2019). Hierarchical flow model-based impact assessment of cyberattacks for critical infrastructures. *IEEE Systems Journal*, *13*(4), 3944–3955. https://doi.org/10.1109/jsyst.2019.2912626

Copyright © 2025 by Author/s and Licensed by JISEM. This is an open access article distributed under the Creative Commons Attribution License which permitsunrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.