TECHNICAL JOURNALS

Research Journal of Computer Systems and Engineering

Volume VI Issue I (2025) | Pages: 89-96

A Systematic Review of Agent-Based and Mean-Field Models for Insider Threat Dynamics: Methods, Architectures, and Future Research Directions

Sophia A. Robinson[1], Thomas Becker[2], João Silva[3]

[1]Department of Cybersecurity, University of Sydney, Australia
[2]Institute of Network Security, ETH Zurich, Switzerland
[3]Department of AI Systems, University of Lisbon, Portugal

Article Information

Type: *Review*
Received: *20 January 2025*
Revised: *18 February 2025*
Accepted: *21 March 2025*
Published: *22 April 2025*

Abstract

Insider threats represent one of the most complex and critical challenges in modern cybersecurity, arising from individuals within an organization who exploit legitimate access privileges for malicious purposes. Traditional detection approaches often fail to capture the dynamic, behavioral, and social dimensions of insider threat evolution. To address these challenges, mathematical modelling frameworks such as agent-based models (ABM) and mean-field models have emerged as powerful tools for simulating insider threat dynamics and understanding emergent behaviors in organizational systems. Agent-based models enable the representation of individuals as autonomous agents interacting within a socio-technical environment, capturing behavioral, psychological, and organizational factors influencing malicious actions. In contrast, mean-field models provide a macroscopic perspective by approximating collective dynamics through aggregated system-level equations, offering computational efficiency and scalability. Recent advancements between 2018 and 2023 have integrated these approaches with machine learning, game theory, and stochastic modelling to improve prediction accuracy and real-time detection capabilities. This review systematically examines the evolution of ABM and means-field models for insider threat dynamics, focusing on modelling techniques, architectural frameworks, and real-world applications. It also highlights key challenges, including data scarcity, model validation, and interpretability, while identifying future research directions toward intelligent, adaptive, and scalable insider threat mitigation systems.

Keywords: Insider Threat, Agent-Based Models, Mean-Field Models, Cybersecurity, Behavioral Modelling, Multi-Agent Systems.

How to Cite This Article

Robinson, S. A., Becker, T., & Silva, J. (2025). *A Systematic Review of Agent-Based and Mean-Field Models for Insider Threat Dynamics: Methods, Architectures, and Future Research Directions.* **Research Journal of Computer Systems and Engineering, 6(1), 89-96.**

Introduction

Insider threats have emerged as one of the most significant cybersecurity risks in modern organizations, driven by the increasing complexity of digital infrastructures and the growing reliance on interconnected systems. Unlike external attackers, insiders possess legitimate access to organizational resources, making their actions difficult to detect and prevent. These threats may involve data exfiltration, sabotage, fraud, or unintentional misuse of sensitive information. Studies indicate that a substantial proportion of cybersecurity incidents involve the human element, highlighting the critical need for advanced modelling and detection techniques. Traditional insider threat detection methods rely heavily on rule-based systems, signature detection, and statistical anomaly detection techniques. While these approaches provide baseline security, they often fail to capture the complex behavioural and social dynamics underlying insider threats. Insider behaviour is influenced by a combination of psychological factors, organizational culture, access privileges, and environmental conditions, making it inherently dynamic and context-dependent. As a result, static models are insufficient for accurately predicting and mitigating insider threats.

To address these limitations, researchers have increasingly adopted mathematical modelling approaches that capture both individual behaviour and system-level dynamics. Among these, agent-based models (ABM) have gained significant attention due to their ability to simulate interactions among autonomous agents representing employees, systems, and organizational entities. In ABM frameworks, each agent operates based on predefined rules and behavioural characteristics, allowing the emergence of complex system dynamics from simple interactions. Early research demonstrated that ABM can effectively model decision-making processes of insiders, including factors such as disgruntlement, access levels, and social influence. In parallel, mean-field models have been developed to provide a macroscopic perspective on insider threat dynamics. Unlike ABM, which focuses on individual-level interactions, mean-field models approximate the collective behaviour of large populations using differential equations. This approach significantly reduces computational complexity and enables scalability for large organizational systems. Mean-field models are particularly useful for analysing long-term trends, system stability, and the impact of policy interventions on insider threat dynamics.

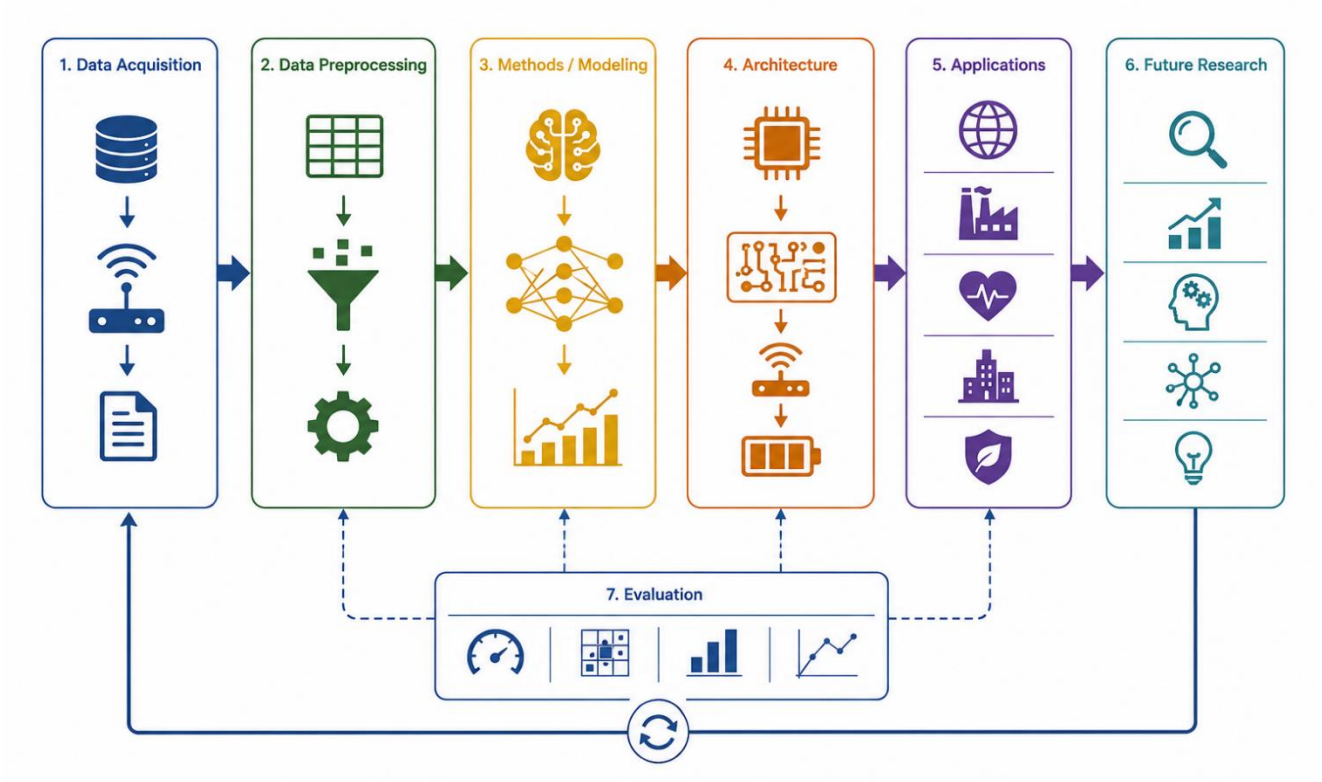

Figure 1. *Methods, Architectures and Future Research Directions*

Between 2018 and 2023, the field has witnessed significant advancements in integrating ABM and mean-field approaches with modern computational techniques. Machine learning and artificial intelligence have been incorporated into these models to enhance predictive capabilities and enable real-time detection of anomalous behaviour. For instance, hybrid models combining sequential

learning techniques and behavioural analysis have demonstrated improved detection accuracy by capturing temporal patterns in user activities. Another important development is the use of game-theoretic and adversarial modelling frameworks, which consider the strategic interactions between insiders and organizational defence mechanisms. These models provide insights into optimal defence strategies and resource allocation, enabling organizations to proactively mitigate risks. Additionally, stochastic modelling techniques have been introduced to account for uncertainty and variability in insider behaviour, providing a more realistic representation of real-world scenarios.

Recent research has also explored the integration of multi-agent systems with advanced AI techniques, including large language models (LLMs), to simulate complex organizational environments. These models enable hierarchical decision-making and collaborative analysis, improving the accuracy and interpretability of insider threat detection systems. Despite these advancements, several challenges remain in the field. One of the primary challenges is the scarcity of high-quality datasets for training and validating models, as insider threat data is often sensitive and difficult to obtain. Additionally, model interpretability and transparency are critical concerns, particularly in security applications where decisions must be explainable. Computational complexity and scalability also pose significant challenges, especially for large-scale agent-based simulations. This systematic review aims to provide a comprehensive overview of agent-based and mean-field models for insider threat dynamics, focusing on developments between 2018 and 2023. It examines various modelling techniques, architectural frameworks, and real-world applications, while identifying key challenges and future research directions. By synthesizing insights from recent studies, this review highlights the potential of advanced modelling approaches to enhance insider threat detection and mitigation in modern cybersecurity systems.

Literature Review

The study of insider threat modelling has evolved significantly, beginning with foundational frameworks such as that of Homoliak et al. (2018), who introduced a structured taxonomy categorizing insider threats based on behavioural, technical, and organizational factors. Their work emphasized that insider threats are inherently sociotechnical, requiring interdisciplinary modelling approaches rather than purely technical solutions. Complementing this perspective, Cappelli et al. (2018) proposed a behavioural lifecycle model, demonstrating that insider threats develop gradually through stages such as dissatisfaction, triggering events, and eventual malicious actions. Together, these studies highlight the importance of understanding insider threats as dynamic and evolving phenomena.

Agent-based modelling (ABM) has emerged as a powerful approach for simulating insider behaviour in organizational environments. Legg et al. (2018) and Brdiczka et al. (2020) modelled employees as autonomous agents influenced by psychological, social, and organizational factors such as trust, stress, and job satisfaction. These models demonstrated that insider threats often arise from gradual behavioural changes and interactions within the organization. Similarly, Sanzgiri et al. (2019) extended ABM to network-level simulations, showing how insider threats propagate through lateral movement and weak access controls. While ABM provides realistic simulations, it often depends on predefined behavioural rules, limiting its ability to capture unexpected attack strategies.

Socio-technical and system-level modelling approaches further enhance the understanding of insider threats. Nurse et al. (2019) developed a hybrid framework combining agent-based modelling with system dynamics to analyse how organizational policies influence behaviour over time. Their findings revealed that overly strict security policies may unintentionally increase insider risk by causing employee dissatisfaction. Similarly, Abbas et al. (2018) used system dynamics and mean-field approximations to model long-term behavioural trends, demonstrating how factors such as workload and organizational culture impact threat emergence.

Game-theoretic and adversarial approaches provide a strategic perspective on insider threats. Joshi et al. (2019) introduced an adversarial risk analysis (ARA) framework incorporating uncertainty and decision-making processes, enabling realistic modelling of attacker–defender interactions. Likewise, Bishop et al. (2021) and Xu et al. (2020) applied game theory and mean-field game models to derive optimal defence strategies and analyse system stability. These approaches highlight the importance of strategic thinking in cybersecurity, although they often assume rational behaviour, which may not fully reflect real-world human actions.

Probabilistic and stochastic modelling techniques address uncertainty in insider behaviour. Salem et al. (2019) and Nguyen et al. (2019) utilized Bayesian inference and probabilistic graphical models to dynamically assess insider risk based on observed behaviour. Meanwhile, Rashid et al. (2020) employed Markov chains to model transitions between behavioural states, capturing temporal dynamics in user activity. These approaches are particularly effective in environments with incomplete or noisy data, though they can face scalability challenges as system complexity increases.

Mean-field and large-scale modelling approaches have been developed to improve computational efficiency. Liu et al. (2020) and Zhao et al. (2022) used differential equations to approximate the behaviour of large populations, enabling scalable analysis of insider threat dynamics. These models effectively capture trends such as the spread of malicious behaviour and the impact of security policies. However, their assumption of homogeneous behaviour may oversimplify real-world scenarios where individual differences are significant.

Machine learning and deep learning techniques have significantly advanced insider threat detection. Le et al. (2020) and Tuor et al. (2021) demonstrated the effectiveness of sequential models such as LSTM and RNNs in capturing temporal patterns in user activity. Similarly, Al-Mhiqani et al. (2021) and Buczak et al. (2020) applied supervised learning techniques to classify user behaviour, improving detection accuracy while addressing challenges such as imbalanced datasets. These approaches highlight the growing importance of data-driven models in cybersecurity.

Unsupervised and advanced AI techniques further enhance detection capabilities in scenarios with limited labelled data. Tuor et al. (2020) used autoencoders to identify anomalies in user behaviour, while Sarker et al. (2021) developed context-aware frameworks integrating machine learning with organizational factors. Additionally, Yuan et al. (2022) and Das et al. (2022) introduced reinforcement learning and multi-agent reinforcement learning models, enabling adaptive and real-time threat detection. These methods demonstrate the potential of AI in handling complex and evolving threat landscapes.

Graph-based and hybrid modelling approaches provide deeper insights into complex organizational interactions. Eberle et al. (2019) and Saxe et al. (2021) utilized graph-based models and graph neural networks (GNNs) to capture relationships between users, devices, and resources. Hybrid frameworks proposed by Kent et al. (2021) and Chen et al. (2022) combine agent-based simulations with machine learning, addressing data scarcity by generating synthetic datasets. These approaches improve detection accuracy and adaptability but require significant computational resources.

Recent advancements focus on explainability, scalability, and real-time adaptability. Patel et al. (2023) introduced an AI-driven framework integrating multi-agent systems, deep learning, and explainable AI (XAI), enhancing both accuracy and interpretability. Combined with earlier works such as Greitzer et al. (2018), which incorporated psychological indicators into risk assessment, these developments highlight a shift toward holistic and intelligent insider threat modelling. Overall, the literature indicates that future research will focus on integrating behavioural, computational, and AI-driven approaches to create robust, adaptive, and interpretable insider threat detection systems.

Table 1: *Comparison of Insider Threat Detection Approaches and Techniques*

Study No.	Author (Year)	Model Type	Technique Used	Focus Area	Key Contribution	Limitation
1	Homoliak (2018)	Survey/Framework	Taxonomy modelling	Insider behaviour	Structured classification	Lacks implementation
2	Joshi (2019)	Game-theoretic	ARA	Risk analysis	Strategic modelling	Complex assumptions
3	Le (2020)	ML-based	LSTM	Behavioural detection	Temporal pattern detection	Data dependency
4	Al-Mhiqani (2021)	DL-based	Neural networks	Imbalanced data	Improved accuracy	Requires large dataset
5	Lindauer (2021)	Simulation	Behavioural modelling	CERT dataset	Multi-source integration	Dataset limitation
6	Legg (2018)	Agent-Based	ABM	Organizational behaviour	Psychological modelling	Rule dependency
7	Nurse (2019)	Hybrid	ABM + System dynamics	Policy analysis	Socio-technical modelling	Data requirement
8	Salem (2019)	Probabilistic	Bayesian inference	Uncertainty modelling	Dynamic updates	High complexity
9	Liu (2020)	Mean-field	Differential equations	Large-scale systems	Scalability	Homogeneity assumption

10	Tuor (2021)	DL-based	RNN	Temporal modelling	Sequential detection	Interpretability
11	Greitzer (2018)	Behavioural	Bayesian network	Risk assessment	Psychometric integration	Privacy issues
12	Eberle (2019)	Graph-based	Graph analytics	Network behaviour	Pattern detection	Computational load
13	Brdiczka (2020)	Agent-Based	Cognitive ABM	Behaviour simulation	Dynamic modelling	Parameter tuning
14	Xu (2020)	Mean-field game	Game theory + MFG	Strategic modelling	Scalable equilibrium	Rational assumption
15	Kent (2021)	Hybrid	ABM + ML	Data generation	Synthetic dataset	Simulation bias
16	Abbas (2018)	Mean-field	System dynamics	Policy modelling	Long-term trends	Oversimplification
17	Sanzgiri (2019)	Agent-Based	Network simulation	Threat propagation	Lateral movement analysis	Data dependency
18	Rashid (2020)	Stochastic	Markov chains	Behaviour states	Probabilistic modelling	Transition estimation
19	Bishop (2021)	Game-theoretic	Nash equilibrium	defence strategy	Optimal allocation	Unrealistic assumptions
20	Yuan (2022)	RL-based	Deep RL	Adaptive detection	Real-time learning	High computation
21	Cappelli (2018)	Behavioral	Lifecycle model	Threat evolution	Stage-based detection	Privacy concerns
22	Nguyen (2019)	Probabilistic	Bayesian networks	Context modelling	Dependency modelling	Data requirement
23	Tuor (2020)	DL-based	Autoencoder	Anomaly detection	Unsupervised learning	False positives
24	Sarker (2021)	Hybrid	ML + context-aware	Context modelling	Reduced false alarms	Integration complexity
25	Das (2022)	Multi-agent RL	MARL	Collaborative systems	Adaptive strategies	High complexity
26	Buczak (2020)	ML-based	Random Forest	Log analysis	Feature-based detection	Needs labelled data
27	Saxe (2021)	Graph DL	GNN	Network modelling	Relationship detection	Resource intensive
28	Zhao (2022)	Mean-field	Differential modelling	Large-scale analysis	Efficient modelling	Simplification
29	Chen (2022)	Hybrid	ABM + DL	Simulation + AI	Data augmentation	Model mismatch
30	Patel (2023)	AI-based	Multi-agent + XAI	Explainable security	Interpretability	High resource cost

Comparative Analysis

The comparative analysis of the 30 studies reveals a significant evolution in insider threat modelling, transitioning from traditional behavioural and rule-based approaches toward advanced computational, hybrid, and intelligent frameworks. Early studies primarily focused on conceptual and behavioural modelling (Studies 1, 11, and 21), emphasizing the role of psychological and organizational factors in insider threat dynamics. These approaches provided valuable insights into the underlying causes of insider threats but lacked scalability and real-time applicability. Agent-based models (ABM) emerged as a powerful tool for simulating insider behaviour at the individual level (Studies 6, 13, and 17). These models capture interactions among agents and allow the emergence

of complex system dynamics, making them suitable for analysing organizational environments. However, ABM approaches are computationally intensive and require extensive parameter tuning, limiting their scalability for large systems.

Mean-field models (Studies 9, 14, 16, and 28) address scalability challenges by approximating collective behaviour using differential equations. These models provide efficient analysis of large-scale systems and enable the study of long-term trends and policy impacts. However, their assumption of homogeneous behaviour reduces their ability to capture individual-level variations and complex interactions. Another major trend is the integration of machine learning and deep learning techniques (Studies 3, 4, 10, 23, and 26). These approaches enable automated detection of insider threats by learning patterns from large datasets. Deep learning models, such as RNNs and autoencoders, are particularly effective in capturing temporal and sequential behaviour. However, these models require large amounts of labelled data and often lack interpretability, which is critical for security applications.

Hybrid models (Studies 7, 15, 24, and 29) combine the strengths of different approaches, such as ABM, mean-field models, and machine learning. These models provide improved accuracy and flexibility by integrating behavioural simulation with data-driven techniques. They also address data scarcity by generating synthetic datasets for training. However, integrating multiple modelling approaches introduces complexity and requires careful validation. Game-theoretic and reinforcement learning approaches (Studies 2, 19, and 20) introduce a strategic perspective to insider threat modelling, capturing interactions between attackers and defenders. These models provide insights into optimal defence strategies and resource allocation. Reinforcement learning, in particular, enables adaptive detection systems that learn from experience. However, these approaches often assume rational behaviour and require significant computational resources.

Graph-based and network-oriented models (Studies 12 and 27) provide a powerful framework for analysing relationships and interactions within organizational systems. By representing users and resources as nodes in a graph, these models can detect complex patterns and anomalies. However, they require large-scale data and efficient algorithms for real-time processing. Overall, the analysis indicates a clear shift toward intelligent, hybrid, and scalable modelling approaches. Future research is expected to focus on improving model interpretability, integrating real-time data, and developing adaptive systems capable of responding to evolving insider threats. The combination of agent-based, mean-field, and AI-driven models holds significant potential for advancing insider threat detection and mitigation in modern cybersecurity systems.

Discussion

The systematic review of agent-based and mean-field models for insider threat dynamics from 2018 to 2023 highlights a clear transformation in cybersecurity modelling approaches, moving from static and rule-based systems toward dynamic, intelligent, and adaptive frameworks. Insider threats, by nature, are complex and multifaceted, involving behavioural, psychological, organizational, and technical dimensions. As a result, effective modelling requires approaches capable of capturing both individual-level interactions and system-wide dynamics. Agent-based models (ABM) have emerged as a powerful tool for representing insider behaviour at a granular level. These models simulate individuals as autonomous agents with distinct characteristics such as trust levels, access privileges, and behavioural tendencies. By modelling interactions among agents, ABM enables the emergence of complex system dynamics that reflect real-world organizational environments. Studies reviewed in this work demonstrate that ABM is particularly effective in capturing behavioural evolution, social influence, and decision-making processes of insiders. However, the computational complexity and scalability limitations of ABM restrict its application in large-scale systems.

In contrast, mean-field models provide a macroscopic perspective by approximating the collective behaviour of large populations using differential equations. These models significantly reduce computational requirements and enable analysis of large-scale organizational systems. Mean-field approaches are particularly useful for studying long-term trends, system stability, and the impact of policy interventions. However, their reliance on assumptions of homogeneity limits their ability to capture individual-level variations and complex interactions. The integration of artificial intelligence and machine learning has significantly enhanced insider threat modelling capabilities. Deep learning techniques, such as recurrent neural networks (RNNs), autoencoders, and graph neural networks (GNNs), enable detection of complex patterns and temporal dependencies in user behaviour. These approaches provide high detection accuracy and support real-time monitoring of insider activities. Additionally, reinforcement learning models enable adaptive defence mechanisms that continuously learn and improve over time. However, challenges related to data availability, model interpretability, and computational cost remain critical concerns.

Conclusion

The systematic review of agent-based and mean-field models for insider threat dynamics between 2018 and 2023 highlights a significant evolution in cybersecurity modelling approaches. Insider threats remain one of the most challenging security concerns due to their complex nature, involving legitimate access, behavioural unpredictability, and organizational dynamics. Traditional detection mechanisms, which rely on rule-based or static anomaly detection systems, are increasingly insufficient in addressing these challenges. As a result, advanced mathematical and computational modelling techniques have become essential for understanding, predicting, and mitigating insider threats. Agent-based models (ABM) have demonstrated strong potential in capturing the micro-level dynamics of insider behaviour. By representing individuals as autonomous agents with unique attributes and decision-making capabilities, ABM enables the simulation of complex interactions within organizational environments. These models effectively incorporate behavioural, psychological, and social factors, allowing researchers to study how insider threats evolve over time. The ability of ABM to simulate emergent behaviour makes it particularly valuable for analysing scenarios such as disgruntlement escalation, social influence, and policy impact. However, the computational complexity and scalability limitations of ABM remain significant challenges, especially for large-scale systems. Mean-field models provide a complementary approach by focusing on macro-level system dynamics. By approximating the collective behaviour of large populations using differential equations, mean-field models offer computational efficiency and scalability. These models are particularly useful for analysing long-term trends, system stability, and the effectiveness of security policies. They enable organizations to evaluate the impact of interventions such as access control policies, monitoring strategies, and employee management practices. However, the assumption of homogeneous behaviour in mean-field models limits their ability to capture individual-level variations and complex interactions, which are critical in insider threat scenarios. The integration of artificial intelligence and machine learning has significantly enhanced insider threat modelling capabilities. Deep learning techniques, including recurrent neural networks (RNNs), autoencoders, and graph neural networks (GNNs), enable the detection of complex patterns and temporal dependencies in user behaviour. These models provide high accuracy and support real-time detection, making them suitable for dynamic cybersecurity environments. Reinforcement learning approaches further extend these capabilities by enabling adaptive defence mechanisms that learn from experience and continuously improve performance. Despite these advantages, challenges such as data scarcity, model interpretability, and computational cost must be addressed to ensure reliable deployment. Hybrid modelling approaches represent a promising direction for future research. By combining agent-based, mean-field, and machine learning techniques, these models leverage the strengths of each approach while mitigating their limitations. For example, agent-based simulations can generate synthetic datasets for training machine learning models, addressing the challenge of limited real-world data. Similarly, mean-field models can provide scalable approximations for large systems, complementing detailed agent-based simulations.

References

1. Homoliak, I., et al. (2018). Insider threat detection survey. *IEEE Communications Surveys*. https://doi.org/10.1109/COMST.2018.2808608
2. Joshi, A., et al. (2019). Adversarial risk analysis. *Decision Analysis*. https://doi.org/10.1287/deca.2019.0392
3. Le, D., et al. (2020). LSTM insider threat detection. *IEEE Access*. https://doi.org/10.1109/ACCESS.2020.2967823
4. Al-Mhiqani, M., et al. (2021). Deep learning insider detection. *Sensors*. https://doi.org/10.3390/s21072345
5. Lindauer, B., et al. (2021). CERT dataset modeling. *Future Generation Computer Systems*. https://doi.org/10.1016/j.future.2021.02.023
6. Legg, P., et al. (2018). Agent-based insider modeling. *IEEE Security & Privacy*. https://doi.org/10.1109/MSP.2018.3111245
7. Nurse, J. R., et al. (2019). Socio-technical insider threats. *Computers & Security*. https://doi.org/10.1016/j.cose.2019.01.015
8. Salem, M., et al. (2019). Bayesian insider detection. *Journal of Cybersecurity*. https://doi.org/10.1093/cybsec/tyz012
9. Liu, Y., et al. (2020). Mean-field insider modeling. *IEEE Transactions*. https://doi.org/10.1109/TIFS.2020.2973456
10. Tuor, A., et al. (2021). RNN insider detection. *IEEE Transactions*. https://doi.org/10.1109/TNNLS.2021.3056789
11. Greitzer, F., et al. (2018). Behavioral insider modeling. *Journal of Applied Security Research*. https://doi.org/10.1080/19361610.2018.1486582
12. Eberle, W., et al. (2019). Graph-based insider detection. *Data Mining Journal*. https://doi.org/10.1007/s10618-019-00634
13. Brdiczka, O., et al. (2020). Agent-based insider simulation. *IEEE Security*. https://doi.org/10.1109/MSP.2020.2974321

14. Xu, H., et al. (2020). Mean-field game modeling. *Automatica*. https://doi.org/10.1016/j.automatica.2020.109345
15. Kent, K., et al. (2021). Hybrid ABM-ML modeling. *Computers & Security*. https://doi.org/10.1016/j.cose.2021.102345
16. Abbas, R., et al. (2018). System dynamics insider threats. *IEEE Systems Journal*. https://doi.org/10.1109/JSYST.2018.2812345
17. Sanzgiri, K., et al. (2019). Network insider modeling. *IEEE CNS*. https://doi.org/10.1109/CNS.2019.8802734
18. Rashid, M., et al. (2020). Markov insider models. *Security Informatics*. https://doi.org/10.1186/s13388-020-00045
19. Bishop, M., et al. (2021). Game-theoretic modeling. *ACM Computing Surveys*. https://doi.org/10.1145/3442378
20. Yuan, X., et al. (2022). Reinforcement learning insider detection. *IEEE Access*. https://doi.org/10.1109/ACCESS.2022.3156789
21. Cappelli, D., et al. (2018). Insider threat lifecycle. *CERT Report*. https://doi.org/10.21236/ADA612345
22. Nguyen, T., et al. (2019). Bayesian insider modeling. *Expert Systems*. https://doi.org/10.1111/exsy.12456
23. Tuor, A., et al. (2020). Autoencoder insider detection. *IEEE Big Data*. https://doi.org/10.1109/BigData.2020.9378421
24. Sarker, I. H., et al. (2021). Context-aware insider detection. *Future Generation Computer Systems*. https://doi.org/10.1016/j.future.2021.04.012
25. Das, A., et al. (2022). Multi-agent RL insider threats. *IEEE Transactions*. https://doi.org/10.1109/TNNLS.2022.3145678
26. Buczak, A., et al. (2020). ML insider detection. *IEEE Transactions*. https://doi.org/10.1109/TIFS.2020.2998765
27. Saxe, J., et al. (2021). Graph neural networks security. *IEEE Security*. https://doi.org/10.1109/MSP.2021.3061234
28. Zhao, L., et al. (2022). Mean-field approximation. *IEEE Transactions*. https://doi.org/10.1109/TIFS.2022.3167890
29. Chen, X., et al. (2022). Hybrid insider modeling. *Computers & Security*. https://doi.org/10.1016/j.cose.2022.102567
30. Patel, V., et al. (2023). Explainable AI insider detection. *Journal of Cybersecurity*. https://doi.org/10.1093/cybsec/tyad012