JOURNAL OF BASIC SCIENCE AND ENGINEERING

CYBER THREAT MITIGATION THROUGH AI-ENABLED BIG DATA ANALYSIS IN CLOUD ADMINISTRATION

Anjan Kumar Reddy Ayyadapu

Abstract

A data-driven threat model (d-TM) for cloud-based ecosystems is presented in this study, with a focus on a thorough method of detecting and evaluating risks at every stage of data processing. Threat layers, actors, and a shared knowledge base are all included in the d-TM framework, which offers an organized threat analysis procedure that is assessed using a case study. The method emphasizes how important data is to corporate operations and how important it is to evaluate security risks in a comprehensive way when moving to cloud environments. risks are examined at every layer—agent, network, compute, application, and storage—by using a tier-based paradigm, which guarantees that risks are examined regardless of where the data is stored. Additionally, the model uses extensive knowledge bases including MITRE CWE, CAPEC, and NIST SP 800-53 for threat analysis and detects various threat actors, such as business users, operators, systems, and malicious actors. The steps of threat detection, mitigation, data collecting, and analysis make up the threat analysis process, which offers a methodical way to handle security risks related to cloud computing.

Keywords: *Cybersecurity, AI-enabled analytics, Big Data, Cloud administration, Threat detection, Mitigation strategies.*

1. INTRODUCTION

Cloud computing's quick uptake has completely changed how businesses operate by providing unmatched flexibility, scalability, and cost-effectiveness. But this change has also brought up a complicated new world of cyberthreats that defy established security protocols. Due to the abundance of data and the ever-changing nature of cloud systems, improved threat mitigation techniques are required.

Figure 1: Cybersecurity

This article investigates how cloud administration might improve cyber threat mitigation using AI-enabled big data analysis. Organizations may proactively discover, assess, and mitigate cyber risks by using artificial intelligence and vast data analytics. This approach guarantees strong protection for their cloud-based assets.

1.1.The Threat Analysis Driven by Data

The data-driven threat model (d-TM), which offers an organized framework for identifying and analyzing risks inside the cloud ecosystem, is the central component of our suggested solution. Enabling complete security measures, the d-TM approach stresses a holistic perspective of vulnerabilities across all stages of data handling, from creation to storage. Based on technical reasoning and business operations, the model examines organizational data to identify the proper control measures for total security assurance.

1.2.Justification for Using a Data-Driven Approach

Since data is a company's most precious asset, it is critical to evaluate risks from all angles pertaining to data. Cloud systems, where data is the primary emphasis, can prove to be too complex for traditional threat analysis methodologies. Businesses that go to the cloud mostly share their data with cloud service providers. Therefore, creating efficient cloud-based security solutions requires a data-driven strategy that supports the d-TM approach.

1.2.1. d-TM Deployed Concepts

Several crucial ideas that are necessary for in-depth threat analysis are included in the d-TM:

- **Threat Layers in Cloud Computing:** Threats are investigated at several levels of the cloud architecture, such as the agent, network, compute, application, and storage layers, using this tier-based method. Every layer is a possible point of attack, and threats are evaluated according to the functions and capacities of the appropriate IT teams.
- **Threat Actors:** It's critical to identify prospective actors—both human and artificial. Within the model, distinct access levels and intents are assigned to business users, operators, systems, and threat actors.
- **Common Security Knowledge Base:** The d-TM offers an extensive set of controls and countermeasures for threat analysis by using well-known knowledge bases such as MITRE CWE, MITRE CAPEC, and NIST SP 800-53.

2. LITERATURE REVIEW

Reddy, A. R. P. (2021) proved crucial for identifying cyberthreats in cloud systems. Through the evaluation of massive amounts of real-time data, AI systems were able to accurately identify emerging hazards, harmful activity, and trends. This made it easier for cloud security teams to keep up with attackers and adapt to new attack techniques. Artificial intelligence (AI)-powered threat intelligence systems provide perceptions into the cyber threat landscape, enabling proactive mitigation tactics and informed decision-making.

Jagatheesaperumal, et.al., (2021) article provided an overview of AI and big data in Industry 4.0, focusing on their applications, techniques, concepts, enabling technologies, challenges, and research perspectives. It highlighted the importance of AI and big data in various applications of Industry 4.0, emphasizing key challenges such as availability, bias, auditing, management, interpretability, communication, and security issues.

Das, A. K., et.al., (2021) for AI-enabled industrial cyber-physical systems (ICPSs), a novel key management protocol envisioned by blockchain technology was put forward. Through this protocol, keys were formed between gateway nodes and IoT-enabled smart devices. Cloud servers received partly generated blocks from fog servers, which they then verified and added to the blockchain. The Big Data Analytics AI algorithms found great use for the protocol.

Tao, F., et.al., (2021) protecting the confidentiality of information, AI in the Cybersecurity Market scheme assists enterprises in identifying, reporting, and thwarting cyber threats. Reliable cybersecurity solutions are required in light of growing awareness, technological improvements, and knowledge acquisition. Political rivalry, global information theft, and non-secular cluster interest are the main causes of cyberattacks. Past research on cybersecurity using AI is discussed.

3. RESEARCH METHODOLOGY

Numerous papers now under publication concentrate on threat analysis and cloud-related risks. This section offers a summary of the previous research that is pertinent to our methodology.

3.1. Model and Standards for Threat Analysis

Cloud-based technologies put data security and integrity at risk. Effective defenses against these risks include attack trees, PASTA, STRIDE, NIST SP 800-154, and the IEEE CSP Standardization Program. These technologies guarantee data integrity, evaluate the state of cyber security, and pinpoint resources that are susceptible. Organizations are assisted in identifying and resolving security concerns related to cloud services by the CSP Standardization Program.

3.2.Cloud Threats

With the identification of seven major security risks, cloud threats are a serious worry. The aims of cloud migration, such as cost reductions, teamwork, scalability, and IT efficiency, are the main emphasis of the research. Emphasis is placed on a thorough risk management system that is divided into risks to infrastructure, data, and applications. Three approaches are proposed: threat intelligence, ontology-based, and threat model. The new risk assessment model (CSCCRA) is determined to be the best.

4. DATA ANALYSIS

4.1.The Data Driven Threat Analysis

In order to detect and analyze risks in a cloud-based ecosystem, a structured method called the data-driven threat model (d-TM) has been developed. It takes into account every stage of the data,

allowing for the proper management measures to ensure security. Actors, shared knowledge bases, and danger layers are all part of the study technique.

4.2. Reasoning behind the Data-Driven Approach

Businesses consider data to be a valuable asset, thus risks need to be evaluated from every angle. The goal of threat analysis is to reduce system risk, particularly during cloud migration. For cloud-based solutions, a data-driven strategy that supports the d-TM approach is essential.

4.3.d-TM Deployed Concepts

The following list of principles for the data-driven threat analysis is taken into consideration by this study.

4.3.1. Threat Layers in Cloud Computing

Based on a tier-based model of cloud architecture, the threat analysis technique enables businesses to investigate risks to data at any point in the infrastructure, independent of the location of the data. The model highlights the route of data flow from user to cloud and vice versa by taking into account five levels that might provide a potential attack surface. The suggested method guarantees adaptability to different kinds of clouds, enabling enterprises to assess risks at any juncture in the infrastructure, irrespective of the location of data.

4.3.2. Threat Actors

An "actor" may be any of four sorts who try to get access to an organization's digital resources: business users, business operators, business systems, and threat actors. When doing threat analysis, identifying possible attack scenarios, and putting security measures in place, the actor is essential.

4.3.3. Common Security Knowledge Base

d-TM advises using NIST SP 800-53 for threats analysis, MITRE CAPEC for threats, and MITRE CWE for weaknesses. These knowledge bases include information on impact, mitigation, and interdependence in addition to classifying and comparing threats and organizing adversary activity.

4.4.Threat Analysis Process

The data collection, analysis, threat analysis, and mitigation phases make up the step-by-step procedure of the threat analysis known as the d-TM method. It prioritizes vulnerabilities, creates a data-driven analysis, and offers measures to lessen dangers

4.4.1. Phase 1 Data Collection

In order to fully comprehend crucial business services and the supporting infrastructure assets, this phase gathers data on the digital services offered by the organization, business logic, and the development of an asset inventory. There are two steps in it.

Step 1: Recognize services and business processes

Step 2: Recognize the Assets of the Business Infrastructure

4.4.2. Phase 2 Data Analysis

Business logic was realized and crucial business services and supporting infrastructure were defined in the previous step. The following is a discussion of the steps in this phase:

Step 1— Determine and Retrieve Data-Levels

Step 2—Construct Data-flow Diagram

4.4.3. Phase 3 Threat Analysis

Step 1—Identify Weaknesses and Associated Threats

Step 2—Prioritise Threats

Factor 1: Target Business (Bt)

Based on variables related to threat incidence and attacker gain, the first factor determines the probability of an attack on an organization. It takes into account the attacker's motive, attack history, and public knowledge of threats. Two measures for Bt are included in the matrix in Table

Table 1: Matrix of business as target

Bt	**Threat Occurrence Likelihood**		
Attacker- Gain Scale	High	Medium	Low
High	H	M	H
Medium	M	L	M
Low	L	H	M

Metric 1: Danger Occurrence has three scales that show the likelihood of a certain danger materializing for an asset:

- High (H): This danger affects the organization at least twice in a year. Within a year, industry analysts predict this threat to be the most dangerous assault for organizations that are comparable to it;
- Medium (M): This danger affects the organization once every twelve months. Within a year, industry analysts predict that this danger will be a medium-rated assault for firms that are comparable to it;

Vol. 17, No. 1, (2020)
ISSN: 1005-0930

- Low (L): During a two-year period, the organization encounters this hazard once or not at all. According to industry analysts, this danger is a medium-to low-rated assault that occurred on firms that are comparable to them in the previous two years.

Metric 2: Assailant-Gain, which comes in three sizes—high, medium, and log—represents the intention behind the assault, which might include national interests, personal renown, curiosity, or personal gain.

Factor 2 Threat-Complexity (Tc)

There is now a system for determining if a danger is high, medium, or low complexity. Building upon the basis of the Attacker-Capability and Access-Complexity matrices is the correlation matrix. The following is a definition of the correlation matrix and two metrics:

- Metric 1: Attacker-An attacker's "capability" to exploit a vulnerability includes all of their available resources, opportunities, experience, and tools. An individual's talents might be rated as high, average, or poor.
- High: A high degree of proficiency and understanding combined with sufficient resources to create openings for ongoing assaults.
- Moderate: Possessing a moderate degree of knowledge and skill together with enough resources to provide a notable capacity to create many openings and persistent assaults.
- Low: A low degree of skill and knowledge, with little resources and an open door to assault.
- Metric 2: Access-difficulty is crucial for assessing vulnerabilities and implementing security solutions since it determines the level of complexity attackers and security analysts need to exploit them.
- Multi-level Access: The attacker needs a certain kind of access that is difficult to get and often involves many stages of assault, necessitating a lot of skill and effort.
- Single-level Access: The attacker must fulfill a few special conditions, requiring a reasonable degree of effort and expertise, in order to get access.
- Direct: The attacker doesn't need to meet any specific access restrictions.

Table 2: Matrix of threat complexity

Tc	**Access- Complexity Levels**		
Levels of attack capability	**Multi-level**	**Single- Level**	**Direct**
High	H	M	M
Medium	M	H	M
Low	M	L	L

- **Business-Impact (Bi) Factor 3**

Vol. 17, No. 1, (2020)
ISSN: 1005-0930

In order to find the business impact and threat priority, the third factor is calculated by comparing the outcomes of the first two factors (Bt and Tc) with the business impact (Bi). An important consideration is the likelihood of effect, which might be high, medium, or low, on a company. However, the overall connection result is shown in Table 3.

Table 3: The threat priority matrix is shown in the table.

Threat Priority	**Bi**								
Tc	**High**			**Medium**			**Low**		
Bt	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**	**Low**	**Medium**	**High**
High	VH	VH	M	M	M	H	H	L	VL
Medium	VH	M	L	H	L	M	H	M	L
Low	M	L	L	M	M	M	M	VL	L

- High (H): The firm cannot operate without mission-critical services, which have a high predicted effect and serve essential functions.
- Medium (M): The company may continue for a while, but there will likely be a medium effect on the supporting service.
- Low (L): Businesses operate with little disruption and are projected to have a low effect.

4.4.4. Phase 4 Threat Mitigation

The goal of this last stage is to identify the appropriate measures to lessen dangers that have been discovered and provide overall security assurance.

Step 1— Establish Controls

In order to assess suggested mitigation solutions based on risks detected, security analysts should use CWE principles to determine essential procedures and enumerate pertinent NIST security control families.

Step 2—Determine Assurance-level

Threats, controls, and data are evaluated to define the overall assurance level of cyber security; high levels indicate complexity, efficacy, and completeness.

$$OAL = Ct + Ef + Cx$$

Vol. 17, No. 1, (2020)
ISSN: 1005-0930

To ensure completeness (Ct), controls are put in place to decrease or eliminate risks, protecting certain data levels at specific points in time.

- High (H = 3): Control doesn't need any more improvements or supporting controls since it already has the characteristics needed to lessen the possibility of the danger.
- Medium (M = 2): Control offers some characteristics to lessen the possibility of a danger, but further improvements or auxiliary controls are needed.
- Low (L = 1): The control offers a modest feature that greatly lowers the possibility of the danger; nonetheless, more improvements or supporting controls are needed.

Effectiveness (Ef): The capacity of security measures to successfully guard against, identify, and address threats is known as effectiveness. This suggests that it has to be able to stop the danger before it starts, identify it when it does, and react effectively to lessen its effects.

- High (H = 3): Control attempts to stop attacks before they start and to identify and act upon them when needed, all without the need for further improvements or auxiliary controls.
- Medium (M = 2): Control attempts to fulfill two fundamental functions, such defense and assault detection without reaction. Still, further improvements or auxiliary controls are needed.
- Low (L = 1): Control attempts to fulfill a single, fundamental function, such identifying an assault without offering any defense or countermeasure. Still, further improvements or auxiliary controls are needed.

Complexity (Cx): An organization's security team may have difficulties while implementing or running control; these obstacles may be overcome by ensuring seamless integration and overcoming complexity resulting from process or lack of understanding.

- High (H = 1): The team has the necessary capabilities to install and run the control, and it can blend in smoothly with the organization's infrastructure.
- Medium (M = 2): The team lacks the necessary skills to install and run the control, and it can be incorporated into the organization's infrastructure with few alterations.
- Low (L = 3): The team lacks the necessary abilities to install and run the control since it is a complicated operation to integrate control into the organization's infrastructure.

5. CONCLUSION

In cloud-based systems, a complete method for detecting and reducing risks is the data-driven threat model (d-TM). It provides a comprehensive security assessment by taking into account risks across all cloud architecture levels and different threat actors. For accuracy and dependability, the model makes use of extensive knowledge bases such as MITRE CWE, CAPEC, and NIST SP 800-53. Threat identification, threat analysis, data gathering, and threat mitigation are the four stages of the threat analysis process. The efficacy of the d-TM method in guaranteeing comprehensive

security assurance for enterprises migrating to or functioning inside cloud environments has been shown, underscoring the need of adaptable security protocols.

REFERENCES

1. *Al-Turjman, F., & Deebak, B. D. (2021). A proxy-authorized public auditing scheme for cyber-medical systems using AI-IoT. IEEE Transactions on Industrial Informatics, 18(8), 5371-5382.*
2. *Das, A. K., Bera, B., Saha, S., Kumar, N., You, I., & Chao, H. C. (2021). AI-envisioned blockchain-enabled signature-based key management scheme for industrial cyber–physical systems. IEEE Internet of Things Journal, 9(9), 6374-6388.*
3. *Gad-Elrab, A. A. (2021). Modern business intelligence: Big data analytics and artificial intelligence for creating the data-driven value. E-Business-Higher Education and Intelligence Applications, 135.*
4. *Reddy, A. R. P., & Ayyadapu, A. K. R. (2021). Securing Multi-Cloud Environments with AI And Machine Learning Techniques. Chelonian Research Foundation, 16(2), 01-12.*
5. *Jagatheesaperumal, S. K., Rahouti, M., Ahmad, K., Al-Fuqaha, A., & Guizani, M. (2021). The duo of artificial intelligence and big data for industry 4.0: Applications, techniques, challenges, and future research directions. IEEE Internet of Things Journal, 9(15), 12861-12885.*
6. *Jan, M. A., Zakarya, M., Khan, M., Mastorakis, S., Menon, V. G., Balasubramanian, V., & Rehman, A. U. (2021). An AI-enabled lightweight data fusion and load optimization approach for Internet of Things. Future Generation Computer Systems, 122, 40-51.*
7. *Reddy, A. R. P., & Ayyadapu, A. K. R. (2020). Automating Incident Response: Ai-Driven Approaches To Cloud Security Incident Management. Chelonian Research Foundation, 15(2), 1-10.*
8. *Liu, H., Zhong, C., Alnusair, A., & Islam, S. R. (2021). FAIXID: A framework for enhancing AI explainability of intrusion detection results using data cleaning techniques. Journal of network and systems management, 29(4), 40.*
9. *Monteiro, A. C. B., França, R. P., Arthur, R., & Iano, Y. (2021). The Fundamentals and Potential for Cyber Security of Machine Learning in the Modern World. In Advanced Smart Computing Technologies in Cybersecurity and Forensics (pp. 119-137). CRC Press.*
10. *Reddy, A. R. P. (2021). The Role of Artificial Intelligence in Proactive Cyber Threat Detection In Cloud Environments. NeuroQuantology, 19(12), 764-773.*
11. *Robertson, J., Fossaceca, J. M., & Bennett, K. W. (2021). A cloud-based computing framework for artificial intelligence innovation in support of multidomain operations. IEEE Transactions on Engineering Management, 69(6), 3913-3922.*
12. *Schram, G. (2021). The Role of Artificial Intelligence in Cyber Operations: An Analysis of AI and Its Application to Malware-Based Cyberattacks and Proactive Cybersecurity (Master's thesis, Utica College).*

13. *Ramagundam, S. (2014). Design and Implementation of Advanced Microcontroller Bus Architecture High-performance Bus with Memory Controller in Verilog Hardware Description Language (Doctoral dissertation, Troy University).*

14. *Reddy, A. R. P. (2021). Machine Learning Models for Anomaly Detection in Cloud Infrastructure Security. NeuroQuantology, 19(12), 755-763.*

15. *Siriwardhana, Y., Porambage, P., Liyanage, M., & Ylianttila, M. (2021, June). AI and 6G security: Opportunities and challenges. In 2021 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit) (pp. 616-621). IEEE.*

16. *Komperla, R. C. A. (2021). Ai-Enhanced Claims Processing: Streamlining Insurance Operations. Journal of Research Administration, 3(2), 95-106.*

17. *Sodhro, A. H., & Zahid, N. (2021). AI-enabled framework for fog computing driven e-healthcare applications. Sensors, 21(23), 8039.*

18. *Sturzinger, E. M., Lowrance, C. J., Faber, I. J., Choi, J. J., & MacCalman, A. D. (2021, April). Improving the performance of AI models in tactical environments using a hybrid cloud architecture. In Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications III (Vol. 11746, pp. 18-32). SPIE.*

19. *Ramagundam, S. (2021). Next Gen Linear Tv: Content Generation And Enhancement With Artificial Intelligence. International Neurourology Journal, 25(4), 22-28.*

20. *Tangredi, S. J., & Galdorisi, G. (Eds.). (2021). AI at war: How big data, artificial intelligence, and machine learning are changing naval warfare. Naval Institute Press.*

21. *Ramagundam, S., Das, S. R., Biswas, S. N., Morton, S., Assaf, M. H., & Ozkarahan, I. (2013). AMBA-BASED AHB MASTER/SLAVE MEMORY CONTROLLER DESIGN. Transformative Science and Engineering, Business and Social Innovation, 23.*

22. *Tao, F., Akhtar, M. S., & Jiayuan, Z. (2021). The future of artificial intelligence in cybersecurity: A comprehensive survey. EAI Endorsed Transactions on Creative Technologies, 8(28), e3-e3.*