**A Multi- Level Clustering Framework for Cybersecurity Risk Stratification in Healthcare: A Dynamic, Overlapping Approach to Threat Classification and Mitigation**

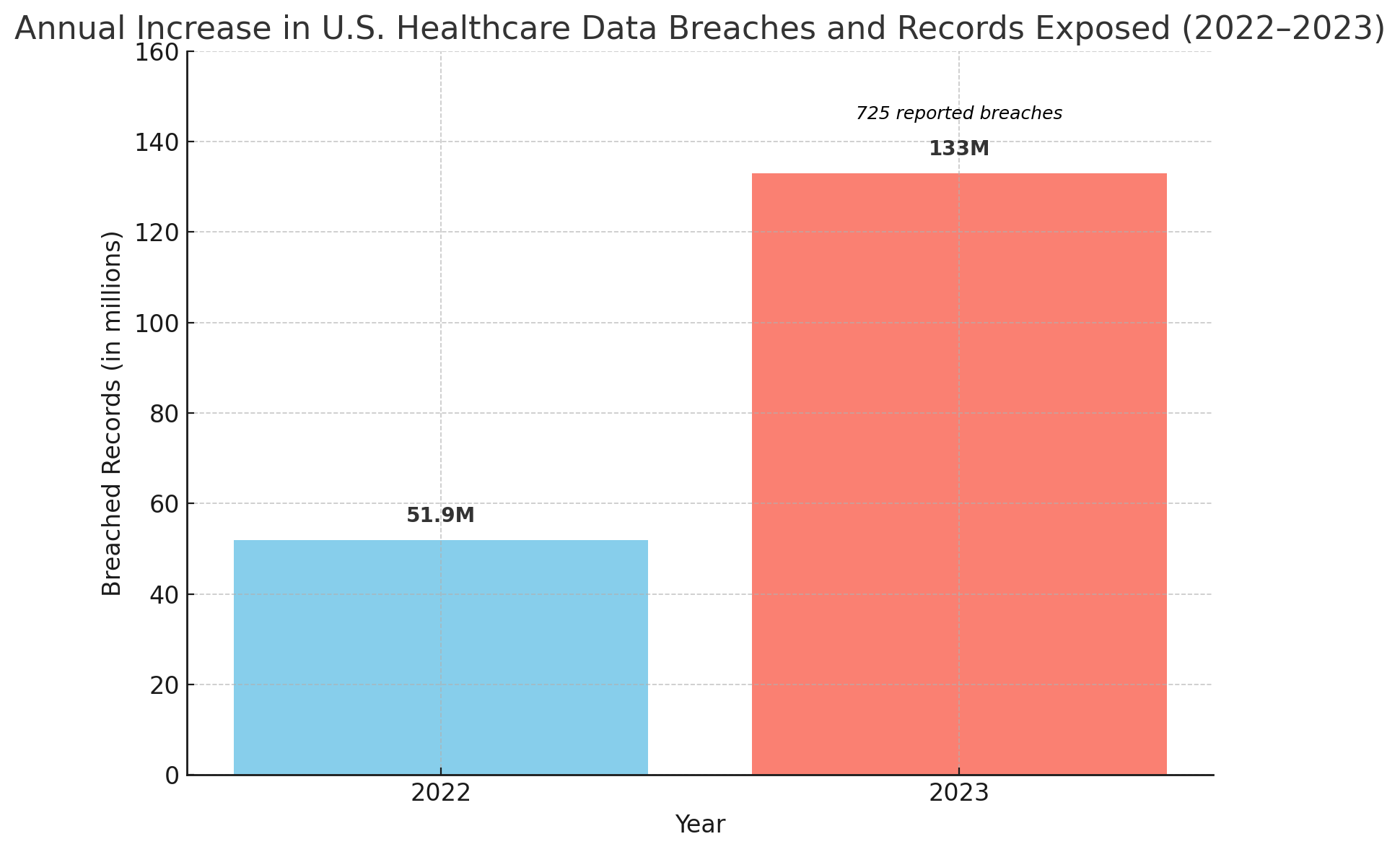
**Abstract**

*The increasing frequency and complexity of cyberattacks targeting the healthcare sector demand innovative approaches to threat classification and mitigation. As healthcare institutions increasingly depend on interconnected digital systems to manage sensitive data, the risk of breaches involving Protected Health Information (PHI) continues to escalate. In response to this challenge, this study proposes a novel, hybrid multi-level clustering framework that integrates Hierarchical Clustering, K-means Clustering, and Fuzzy C-means (FCM) Clustering to dynamically stratify cybersecurity threats in the U.S. healthcare sector. Utilizing a diverse dataset comprising over 1,200 breach incidents from the HHS Breach Portal, enriched with threat intelligence feeds and simulated SIEM logs, the model effectively captures evolving threats based on severity, frequency, and financial impact. Unlike previous models, this framework supports partial membership handling and real-time threat assessment, significantly improving threat categorization and predictive capabilities. Results demonstrate superior performance compared to traditional K-means clustering, with improved accuracy, coherence, and adaptability. Evaluation metrics confirm the model’s efficacy in enhancing decision-making, resource prioritization, and compliance adherence. This approach offers practical applications for healthcare institutions aiming to fortify digital infrastructure against sophisticated, evolving threats. Recommendations include adopting this hybrid model for proactive threat detection, integrating real-time data inputs, and promoting further research into dynamic, overlapping clustering methodologies. The findings present a valuable tool for researchers, policymakers, and practitioners striving to improve cybersecurity resilience and regulatory compliance within the healthcare industry.*

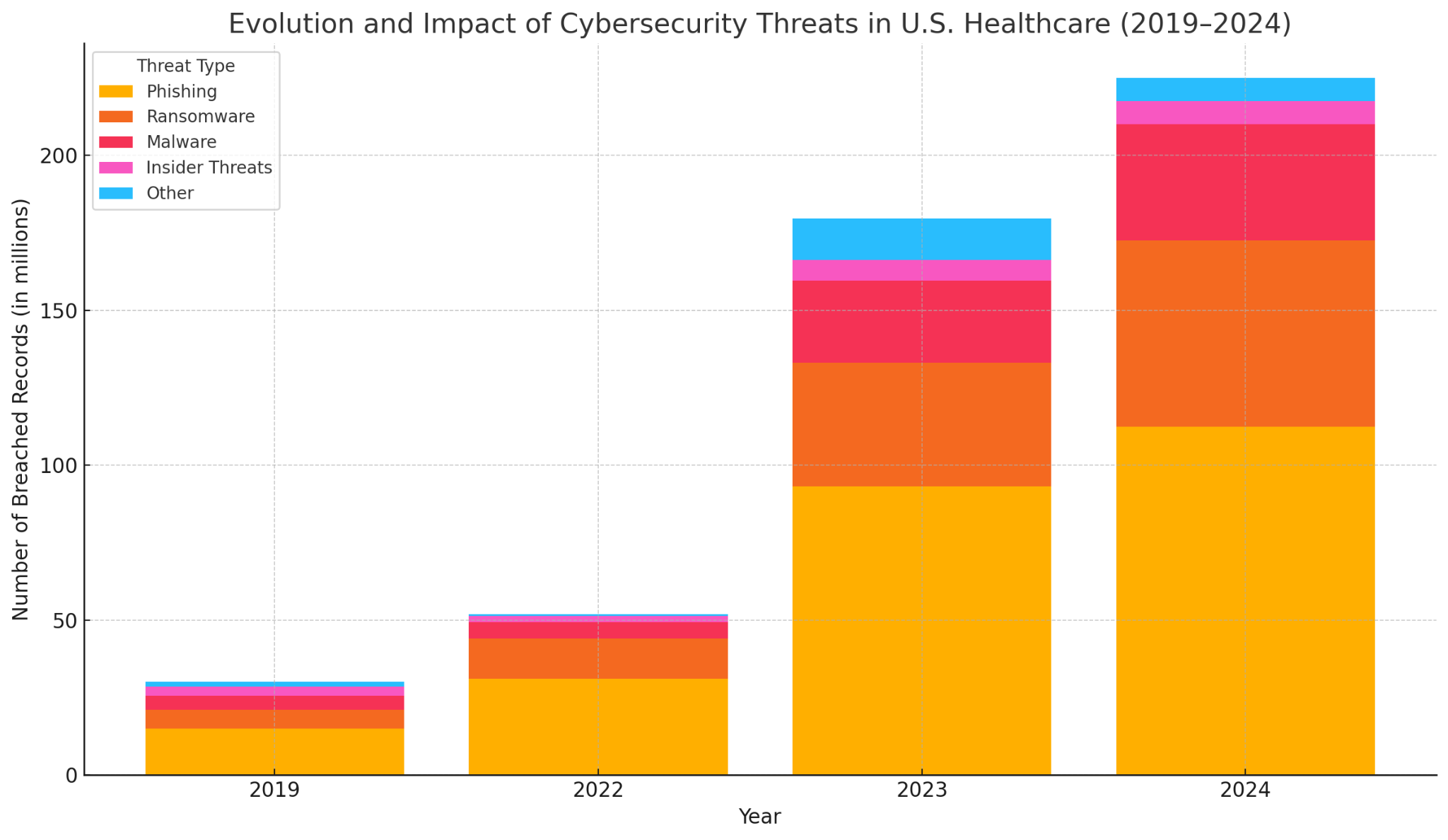
**Keywords: Cybersecurity, Hybrid Clustering, Healthcare Sector, Fuzzy C-means Clustering, Threat Stratification.**

**1. I5rdzntroduction**

Cybersecurity within the U.S. healthcare sector has become a pressing concern in the 21st century, owing to the sector’s custodianship of vast volumes of sensitive data, particularly Protected Health Information (PHI), which holds considerable monetary value on illicit markets. The attractiveness of healthcare institutions as cybercrime targets has led to a marked rise in both the frequency and severity of attacks, resulting in operational disruptions, financial liabilities, and the compromise of patient confidentiality. According to Alder (2024), data from the U.S. Department of Health and Human Services (HHS) indicated that 2023 witnessed 725 reported data breaches, affecting over 133 million patient records—more than double the 51.9 million records compromised in 2022. This surge demonstrates a systemic escalation of threats, attributable not to isolated lapses but to structural vulnerabilities embedded across the healthcare ecosystem.  
This sharp rise is visually captured in Figure 1, which shows the year-on-year increase in healthcare data breaches and the number of compromised patient records between 2022 and 2023.



**Figure 1:** **Annual Increase in Healthcare Data Breaches and Records Exposed (2022–2023).**

George et al. (2024) posit that such persistent exposure arises from a convergence of outdated technological infrastructure, inconsistent regulatory compliance, and increasingly sophisticated threat vectors. Notably, hacking and IT incidents accounted for 80% of reported breaches in 2023, compared to 49% in 2019, illustrating the growing prevalence of technical intrusions (McKeon, 2023). The financial repercussions are equally significant; the average cost of a healthcare data breach climbed to $10.93 million in 2023, while the per-record cost reached $408, significantly surpassing the cross-industry average of $148 (Alder, 2023). These figures underscore the need for more effective cybersecurity frameworks and the strategic allocation of protective resources.  
Figure 2 illustrates the evolving composition and scale of cybersecurity threats in the U.S. healthcare sector between 2019 and 2024, emphasizing the increasing prominence of phishing and ransomware over time.  


**Figure 2: Evolution and Impact of Cybersecurity Threats in U.S. Healthcare (2019–2024)**

The study by Balogun (2025), “Strengthening Compliance with Data Privacy Regulations in U.S. Healthcare Cybersecurity,” provided a comprehensive evaluation of current compliance strategies and cybersecurity practices. By employing K-means clustering and multivariate regression analysis, the study identified three primary threat categories: moderate-risk threats, such as phishing and ransomware; high-impact threats, including insider breaches; and persistent, high-severity threats, primarily malware. A key finding of this research—the statistically significant relationship between financial investment and cybersecurity compliance—reinforced the notion that budget allocation directly influences an institution’s capability to protect electronic health records, medical IoT systems, and core IT infrastructure.

Nonetheless, upon critical analysis, it was observed that methodological limitations constrained the study. Specifically, a static clustering model assumed discrete, non-overlapping threat categories, neglecting the complex interplay among attack types. For instance, Cheng and Wang (2022) explains that phishing campaigns often serve as precursors to ransomware deployment, which may subsequently evolve into system-wide malware attacks—demonstrating how individual incidents frequently span multiple classifications. Moreover, the exclusion of real-time threat intelligence, such as Security Information and Event Management (SIEM) data or Intrusion Detection Systems (IDS), restricted the study's capacity for dynamic risk assessment.

Addressing these limitations, the present study proposes a Multi-Level Clustering Framework for Cybersecurity Risk Stratification in Healthcare. This integrative model employs hierarchical clustering, K-means, and Fuzzy C-Means (FCM) algorithms to reflect contemporary cyber threats' overlapping and adaptive attributes. In contrast to rigid classification systems, Iglesias et al. (2019) note that fuzzy clustering facilitates partial membership across threat categories, thereby capturing the fluidity and multidimensionality of cyberattacks. The 2024 Change Healthcare breach exemplifies this complexity, involving credential theft, ransomware, and lateral movement, each implicating multiple clusters simultaneously.

This study incorporates both empirical and simulated datasets, including breach reports from the HHS, publicly accessible threat intelligence feeds, and synthetically generated SIEM logs. Sharma and Bantan (2024) argue that this multifaceted data integration supports a realistic, continuously updating risk stratification framework capable of identifying and categorizing emerging threat vectors. Real-world case studies serve as validation mechanisms; for example, the Universal Health Services (UHS) ransomware incident in 2020, triggered by a phishing entry point, led to $67 million in recovery expenses and nationwide operational disruption (Lyngaas, 2021). Similarly, the Scripps Health breach in 2021 combined insider activity with malware dissemination, underscoring the need for nonlinear threat modeling (Alder, 2021).

According to Dimakopoulou and Rantos (2024), the necessity for dynamic models is further substantiated by industry-wide statistics indicating an escalation in both threat volume and sophistication. In 2024, 67% of global healthcare organizations reported ransomware incidents, an increase from 34% in 2021, while phishing remained the leading attack vector, involved in over 90% of reported breaches (Mahendru, 2024). Human-related vulnerabilities, including inadequate risk awareness and the continued use of obsolete IT systems, further exacerbate exposure. Sar (2024) adds that aging equipment and third-party service providers serve as recurring entry points, with external vendors responsible for 58% of breaches in 2023.

Existing literature affirms the value of advanced analytical methodologies in cybersecurity (Hossain & Islam, 2024; Sarker, 2022). Alfoudi et al. (2022) assert that clustering algorithms are widely employed within Intrusion Detection Systems to improve threat identification accuracy. Additionally, emerging research into quantum machine learning models indicates their potential for enhancing vulnerability classification. Applications of Fuzzy Cognitive Mapping (FCM) in telehealth security further exemplify the viability of modeling interrelated threats for real-time decision-making (Poleto et al., 2021). These innovations align with the current study's objectives, which aim to improve the specificity and responsiveness of cybersecurity protocols through enhanced clustering strategies.

Building on the foundational insights offered by Balogun, this research aims to resolve earlier methodological limitations by incorporating dynamic threat modeling and allowing for cross-cluster threat representation. Iglesias et al. (2019) emphasize that the framework reinforces the critical role of financial investment in cybersecurity resilience and facilitates targeted allocation of protective resources. Through multidimensional clustering and integration of live threat intelligence, the model is designed to improve institutional decision-making and foster greater regulatory alignment with standards such as HIPAA, HITECH, and NIST (Elendu et al., 2024). This study provides a critical advancement in cybersecurity risk stratification for the healthcare sector by proposing a dynamic, multi-level clustering model. Unlike static models, the hybrid framework integrates Hierarchical Clustering, K-means, and Fuzzy C-means algorithms to detect, classify, and mitigate evolving threats more effectively. Its novel approach addresses overlapping threat categories, enhancing real-time risk assessment, decision-making, and regulatory compliance. This research offers valuable insights for academia, industry practitioners, and policymakers. This study aims to propose and evaluate a dynamic, multi-level clustering framework for stratifying cybersecurity risks in the U.S. healthcare sector by applying integrated hierarchical, K-means hybrid, and fuzzy clustering techniques using real-world and simulated cybersecurity data, by achieving the following objectives:

1. Formulate a hybrid multi-level clustering model that uses hierarchical clustering, K-means, and fuzzy C-means algorithms for classifying cybersecurity threats based on their severity, frequency, and financial impact.
2. Integrate and analyze historical and simulated real-time cybersecurity datasets to reflect the evolving and overlapping nature of cyber threats in the healthcare sector.
3. Assess the performance of the proposed clustering model using established clustering and classification metrics and benchmark it against traditional K-means approaches.
4. Critically examine how the improved stratification of cybersecurity threats can inform risk-based decision-making and resource prioritization in healthcare cybersecurity compliance and strategy.

**2. Literature Review**

Conventional cybersecurity risk classification models have long functioned as foundational instruments for threat identification and management, particularly in data-intensive sectors such as healthcare (Ameedeen et al., 2024; Ajayi et al., 2025). These models frequently employ static methodologies, with K-means clustering emerging as one of the most prevalent techniques. K-means classifies data into a predetermined number of discrete, non-overlapping clusters based on similarity metrics, thereby enabling organizations to categorize cyber threats by attributes such as severity, recurrence, and economic implications (Miraftabzadeh et al., 2023; Balogun, 2025). According to Balogun (2025), this method was applied to the U.S. healthcare context and delineated three distinct clusters: moderate-risk threats such as phishing and ransomware, high-impact insider threats, and severe, persistent threats primarily involving malware. AlDaajeh and Alrabaee (2024) posit that this structured taxonomy facilitated a clearer understanding of the threat environment and revealed the correlation between cybersecurity investment and regulatory compliance.

The utility of K-means and comparable static classification methods has extended to multiple cybersecurity domains, including intrusion detection, malware analysis, and anomaly recognition (Begum et al., 2024; Kolade et al., 2025). Sufi and Alsulami (2025) argue that these clustering techniques assist in identifying recurrent attack patterns, enabling institutions to prioritize countermeasures and allocate resources more efficiently. However, their effectiveness diminishes within today’s rapidly evolving threat environment, where attacks often exhibit multidimensional characteristics and defy rigid taxonomic structures.

A primary limitation of static models lies in their requirement for mutually exclusive cluster membership, which inadequately represents real-world cyber incidents' overlapping and interconnected nature (Zhu et al., 2024; Obioha-Val, 2025). Gómez-Hernández et al. (2023) contend that threats rarely exist in isolation; a single phishing campaign, for example, may escalate into ransomware deployment and culminate in large-scale data exfiltration, thus straddling multiple risk categories. The fixed architecture of static models often reduces the complexity of such scenarios into oversimplified classifications, thereby undermining the accuracy of risk assessments (Wani, 2024; Olutimehin, 2025).

Moreover, Li et al. (2024) note that static models are inherently retrospective, relying solely on historical data and lacking mechanisms to adapt to novel or emergent threat vectors. Their exclusion of real-time threat intelligence—such as Security Information and Event Management (SIEM) outputs and intrusion alerts—further impairs their responsiveness and operational relevance (González-Granadillo et al., 2021; Balogun et al., 2025). In a healthcare context, where cyber threats are escalating in sophistication and scope, such deficiencies render static frameworks insufficient. These limitations underscore the necessity for adaptive, multi-tiered classification systems that integrate dynamic data inputs and permit partial cluster membership, thereby offering a more nuanced and responsive model for situational awareness and strategic mitigation.

**Clustering Techniques in Cybersecurity Research**

Clustering algorithms have garnered substantial attention in cybersecurity research due to their ability to analyze complex datasets and classify diverse threat types, particularly within data-intensive sectors like healthcare (Razaque et al., 2025; Balogun et al., 2025). According to Eskandari et al. (2022), hierarchical clustering, K-means clustering, and Fuzzy C-Means (FCM) clustering represent methodologically distinct yet complementary approaches, each contributing specific advantages to cybersecurity risk stratification.

Hierarchical clustering operates by generating a nested cluster structure using either agglomerative or divisive logic. Agglomerative methods initiate from individual data points, successively merging based on similarity, whereas divisive approaches begin with a cluster incrementally partitioned cluster (Wani, 2024; Obioha-Val et al., 2025). The resulting dendrogram visually depicts data relationships, proving especially useful in intrusion detection systems that require layered insight into network anomalies (Azam et al., 2023; Olutimehin, 2025). Kumar et al. (2024) note that one key strength of hierarchical clustering lies in its ability to determine optimal cluster counts without relying on predefined parameters. However, its computational complexity poses a scalability issue when applied to large, real-time cyber security datasets.

In contrast, K-means clustering is frequently adopted for its computational efficiency and algorithmic simplicity (Ikotun et al., 2021; Obioha-Val et al., 2025). This technique partitions data into a set number of clusters by minimizing intra-cluster variance, a process particularly suited to categorizing recurring breach patterns. Aziz and Bestak (2024) posit that K-means has demonstrated practical efficacy in cyber threat detection, supporting prioritization efforts and resource allocation. Balogun (2025), for instance, employed K-means to segment threats in the U.S. healthcare system into moderate, high-impact, and severe categories, thus enabling a structured strategic response. Nonetheless, Sun and Huang (2024) argue that the assumption of mutually exclusive cluster membership inherent in K-means limits its ability to capture the interrelatedness of real-world threats. Hybrid approaches combining K-means with hierarchical clustering have been introduced to address this limitation by improving boundary delineation in complex datasets.

Fuzzy C-Means clustering offers an alternative that incorporates partial cluster membership, reflecting the ambiguity and overlap commonly found in cyber incidents. Grounded in fuzzy logic, FCM allows each data point to belong to multiple clusters with varying degrees of membership (Hussein et al., 2023; Olutimehin, 2025). Butt et al. (2021) avers that this method has proven effective in modeling telehealth security systems and in analyzing multifaceted vulnerabilities through tools like fuzzy TOPSIS and Fuzzy Cognitive Mapping. By enabling the flexible categorization of compound threats, such as phishing attacks that culminate in ransomware deployment, FCM addresses a critical gap in traditional models. An integrated framework that combines hierarchical, K-means, and fuzzy clustering methodologies offers a more robust and adaptive approach to threat classification in complex, evolving environments.

**Dynamic Risk Stratification and Real-Time Data Integration**

In the context of an increasingly volatile healthcare cybersecurity environment, the integration of real-time data has become indispensable for effective risk stratification and timely threat mitigation (Aljohani, 2023; Obioha-Val et al., 2025). According to González-Granadillo et al. (2021), Security Information and Event Management (SIEM) systems, Intrusion Detection Systems (IDS), and external threat intelligence feeds constitute the foundational tools supporting this dynamic model. SIEM platforms, as defined by the National Institute of Standards and Technology (NIST), consolidate data across disparate security components to generate centralized, actionable insights. Concurrently, IDS monitors network traffic to detect anomalies, while threat intelligence feeds provide current information on vulnerabilities and emergent attack tactics. These technologies enhance situational awareness and enable organizations to detect, interpret, and respond to threats with minimal delay.

The limitations of static classification models are evident through high-profile cyber incidents that exposed such systems' rigidity and retrospective nature. Özeren (2025) observes that the 2024 Change Healthcare ransomware attack, executed by the ALPHV/BlackCat group, disrupted billing operations across the U.S. and compromised data belonging to over 100 million patients. Similarly, the 2021 Scripps Health breach and the 2020 Universal Health Services (UHS) ransomware attack incurred severe operational and financial consequences (Alder, 2021; Olutimehin et al., 2025). Rihan et al. (2023) contend that these events underscore the inadequacy of traditional frameworks that rely on fixed threat signatures and static patterns, making them ill-suited to counter dynamic and evolving attack strategies.

Unlike static models, dynamic stratification systems utilize streaming data to enhance responsiveness and analytical precision. Ali et al. (2024) posit that this adaptability is essential for confronting cybercriminal tactics that evolve in real time. Megherbi et al. (2024) introduced StreamSpot, a framework that leverages heterogeneous data streams for anomaly detection, processing each event upon arrival to uncover malicious behavior. Similarly, Zipperle et al. (2022) developed SAQL, a stream-based query system capable of identifying behavioral deviations through continuous event stream analysis. These models highlight the strategic advantage of incorporating real-time analytics into cybersecurity architecture.

Furthermore, Abdelkader et al. (2024) argue that integrating real-time and simulated datasets strengthens the robustness of cyber defense mechanisms. While real-time inputs facilitate immediate detection of novel exploits, simulated data allow for testing under controlled, hypothetical scenarios. This dual strategy supports comprehensive threat preparedness and ensures that healthcare institutions can proactively defend against complex and fast-evolving cyber threats, thus safeguarding critical infrastructure and sensitive patient information (Sun et al., 2023; Salako et al., 2024).

**Overlapping Threat Categories and Hybrid Threat Events**

The escalating complexity of cyberattacks poses substantial challenges to conventional cybersecurity frameworks, particularly within the healthcare sector, where breaches have far-reaching operational, financial, and ethical implications (Tariq, 2024; Alao et al., 2024). According to Steingartner et al. (2021), contemporary threats frequently appear as hybrid events, combining multiple vectors in a single incident. A typical scenario involves phishing campaigns serving as the preliminary access point for ransomware deployment (Tan et al., 2024; Gbadebo et al., 2024). These attacks often deceive users into divulging credentials or executing malicious payloads, creating avenues for deeper infiltration. The 2020 Universal Health Services (UHS) breach exemplifies this tactic, as a phishing email triggered a widespread ransomware infection that incapacitated over 400 healthcare facilities and caused estimated financial losses of $67 million (Lyngaas, 2021).

Temara (2024) notes that the confluence of credential theft and malware execution is increasingly characteristic of modern cyber intrusions. The 2024 Change Healthcare ransomware attack, attributed to the ALPHV/BlackCat group, illustrates this strategy (Özeren, 2025). Attackers exploited compromised login credentials to gain unauthorized access, subsequently deploying ransomware that exposed personal data belonging to over 100 million individuals. UnitedHealth Group reportedly paid $22 million in ransom in response (Lyngaas, 2021). Similarly, the 2021 Scripps Health breach involved unauthorized system access followed by malware propagation and data exfiltration, impacting approximately 1.2 million patients (Alder, 2021). The disruptions and the $3.5 million legal settlement reflect the interplay between technical flaws and human vulnerabilities in healthcare cybersecurity incidents.

These examples reveal a recurring pattern in which cyber threats transcend discrete classifications and instead manifest as interwoven events. Traditional risk stratification models—such as those based on K-means clustering—presume mutually exclusive threat categories and, therefore, fall short in capturing the intertwined nature of modern cyber events (Byrapuneni & Saidireddy, 2024; Joseph, 2024). Al-Sabbagh et al. (2024) argue that such models fail to accommodate the concurrency of phishing, ransomware, and insider threats, which frequently coalesce within a single breach.

To overcome this methodological shortcoming, the proposed study advocates for a multi-level clustering framework that integrates hierarchical clustering, K-means, and Fuzzy C-Means (FCM). Hierarchical clustering reveals embedded threat structures; K-means provides computational efficiency in grouping similar events; and FCM facilitates partial membership across clusters, thus accurately modeling the overlap inherent in hybrid threats (Beiranvand et al., 2024; Kolade et al., 2024). Javadpour et al. (2023) posit that this composite approach delivers a flexible, scalable, and contextually appropriate model for addressing healthcare institutions' multifaceted cybersecurity challenges.

**Compliance, Resource Allocation, and Decision-Making in Healthcare Cybersecurity**

Compliance with data privacy regulations such as the Health Insurance Portability and Accountability Act (HIPAA) and the Health Information Technology for Economic and Clinical Health (HITECH) Act remains an institutional imperative within the U.S. healthcare system (Isibor, 2024; Mayeke et al., 2024). However, Isibor (2024) observes that adherence levels remain insufficient. Balogun (2025) reports that only 38% of healthcare providers demonstrate full compliance with the HIPAA Security Rule, indicating a significant regulatory gap. This deficiency has direct financial repercussions, as penalties for non-compliance increased from an average of $529,000 in 2020 to $863,000 by 2023 (SAI360, 2023). Additionally, Zorabedian (2024) places the global average cost of a breach at $4.88 million, underscoring the growing fiscal risk of cybersecurity failures.

According to Ilori et al. (2024), financial resource allocation plays a central role in shaping organizational compliance behavior. Balogun (2025) identifies budgetary commitment as the only statistically significant predictor of cybersecurity compliance (p = 0.0178), revealing a direct correlation between institutional investment and regulatory conformance. Despite this evidence, Pentera (2024) reveals that 51% of surveyed enterprises experienced a breach within two years, and only 7% reported no damage. Pentera (2024) contends that these findings reflect persistent underinvestment and stagnation in cybersecurity funding across many healthcare organizations, reinforcing systemic vulnerabilities.

To support strategic decision-making, Bhol (2025) argues that Multi-Criteria Decision-Making (MCDM) methodologies provide structured mechanisms for prioritizing cybersecurity interventions. Madanchian and Taherdoost (2023) applied the Fuzzy Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) to evaluate security requirements in healthcare software development, enabling decision-making under uncertain and variable conditions. Complementarily, Balogun (2025) introduced the Cybersecurity Maturity Model (CMM) as a procedural framework aimed at guiding incremental security improvements and informing resource distribution strategies.

Nevertheless, Chhabra Roy and Prabhakaran (2022) posit that while these models offer structure, they often lack integration with real-time risk data, limiting their responsiveness to rapidly evolving threats. Addressing this shortfall, the present study introduces a comprehensive framework that interlinks budget allocation, compliance performance, and dynamic threat monitoring. By combining real-time risk stratification with financial and regulatory metrics, this model enhances decision-making precision and enables healthcare institutions to respond effectively to the increasing complexity and pace of cybersecurity threats.

### **3. Methodology**

This study employs a quantitative approach using hierarchical, K-means, and fuzzy C-means (FCM) clustering for cybersecurity risk stratification. The objective is to classify threats based on severity, frequency, and financial impact to enhance decision-making and compliance strategies. Data was sourced from the HHS Breach Portal and MITRE ATT&CK Framework. The dataset includes attributes such as affected records, breach type, cause, and financial impact.

Data was cleaned, normalized, and standardized using Z-score normalization:

Where:

* X = raw value, μ = mean, σ = standard deviation.

Categorical variables were encoded using one-hot encoding. The proposed multi-level framework integrates Hierarchical Clustering, K-means Clustering, and Fuzzy C-means (FCM) clustering. Agglomerative clustering was applied to establish the optimal number of clusters using Ward’s linkage criterion:

Where:

* Dij = distance between clusters iii and jjj,
* ni,nj​ = sizes of clusters i and j,
* ci,cj​ = centroids of clusters i and j.

The optimal number of clusters (k) was determined using the Elbow Method, where the Within-Cluster Sum of Squares (WCSS) is calculated as:

The K-means algorithm partitions data into k clusters identified by hierarchical clustering. The objective function is:

Where:

* J = objective function,
* xj ​ = data point,
* μi​ = centroid of cluster iii.

K-means++ initialization was employed for optimal centroid selection. Convergence was achieved when changes in centroid positions were below 10−6.

FCM allows overlapping membership across clusters, calculated by:

Where:

* m = fuzziness coefficient (typically 2),
* c = number of clusters.

The centroid calculation is:

The Fuzzy Partition Coefficient (FPC), used to evaluate clustering quality, is:

Higher FPC values indicate well-defined clusters. Also, clustering performance was evaluated using:

1. Silhouette Score:

Where:

* a = average intra-cluster distance,
* b = minimum average inter-cluster distance.

1. Davies-Bouldin Index (DBI):

Where:

* σi = average distances within clusters iii and j,
* d(ci,cj) = distance between centroids ci and cj​.

1. Fuzzy Partition Coefficient (FPC): This is already defined above.

The proposed framework was benchmarked against traditional K-means clustering. Risk scores were computed using a weighted formula integrating severity (SSS), frequency (F), and financial impact (I):

Where:

* α, β,γ = weighting coefficients derived from the Analytical Hierarchy Process (AHP).

**4. Results**

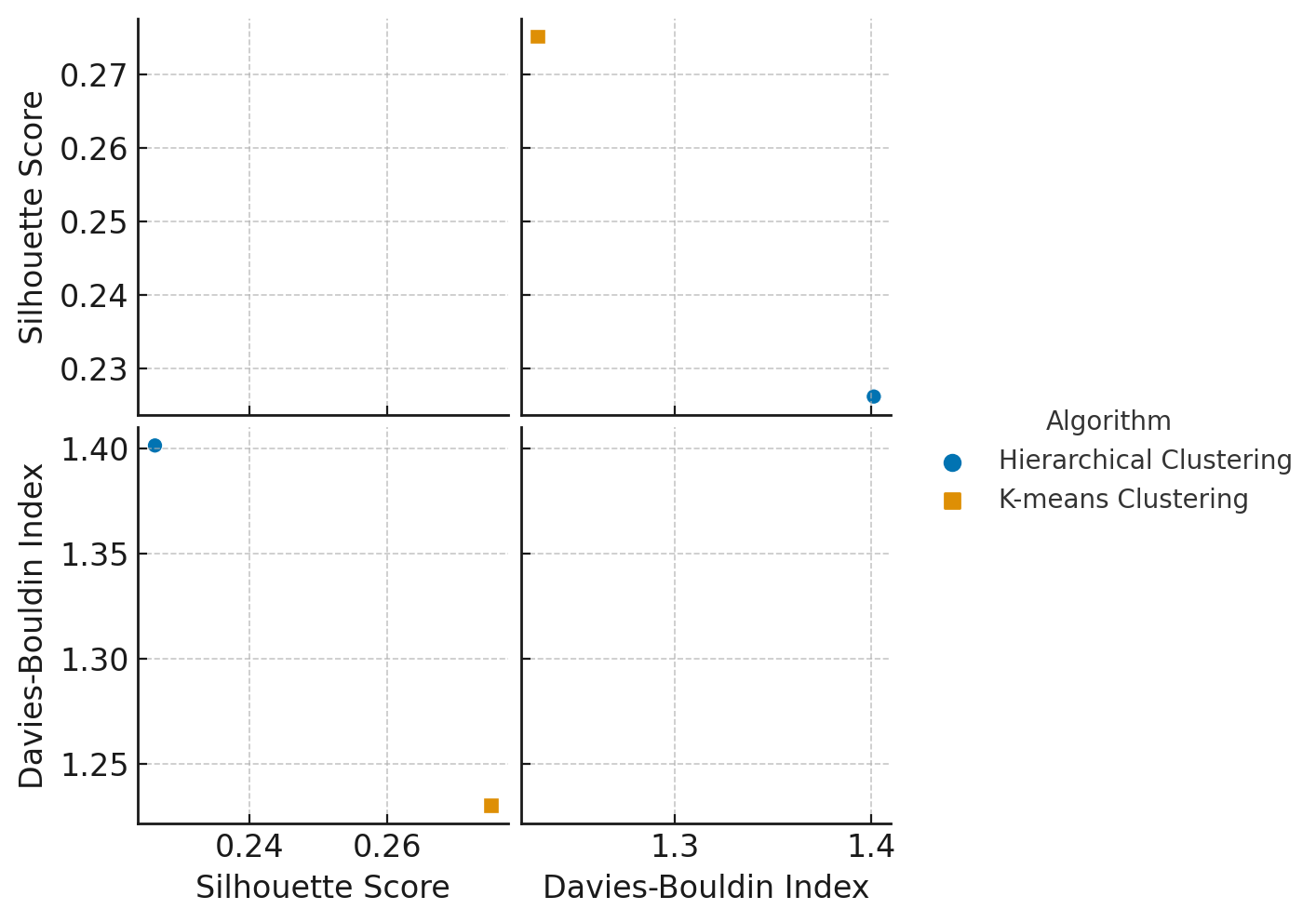
The results of the clustering model are evaluated using relevant metrics to assess performance and efficacy, as explained in the methods above. The performance of the clustering algorithms was assessed using the Silhouette Score, Davies-Bouldin Index, and Fuzzy Partition Coefficient (FPC). The results are presented in Table 1.

#### **Table 1: Clustering Model Evaluation Results**

|  |  |  |
| --- | --- | --- |
| **Algorithm** | **Silhouette Score** | **Davies-Bouldin Index** |
| Hierarchical Clustering | 0.226 | 1.402 |
| K-means Clustering | 0.275 | 1.230 |

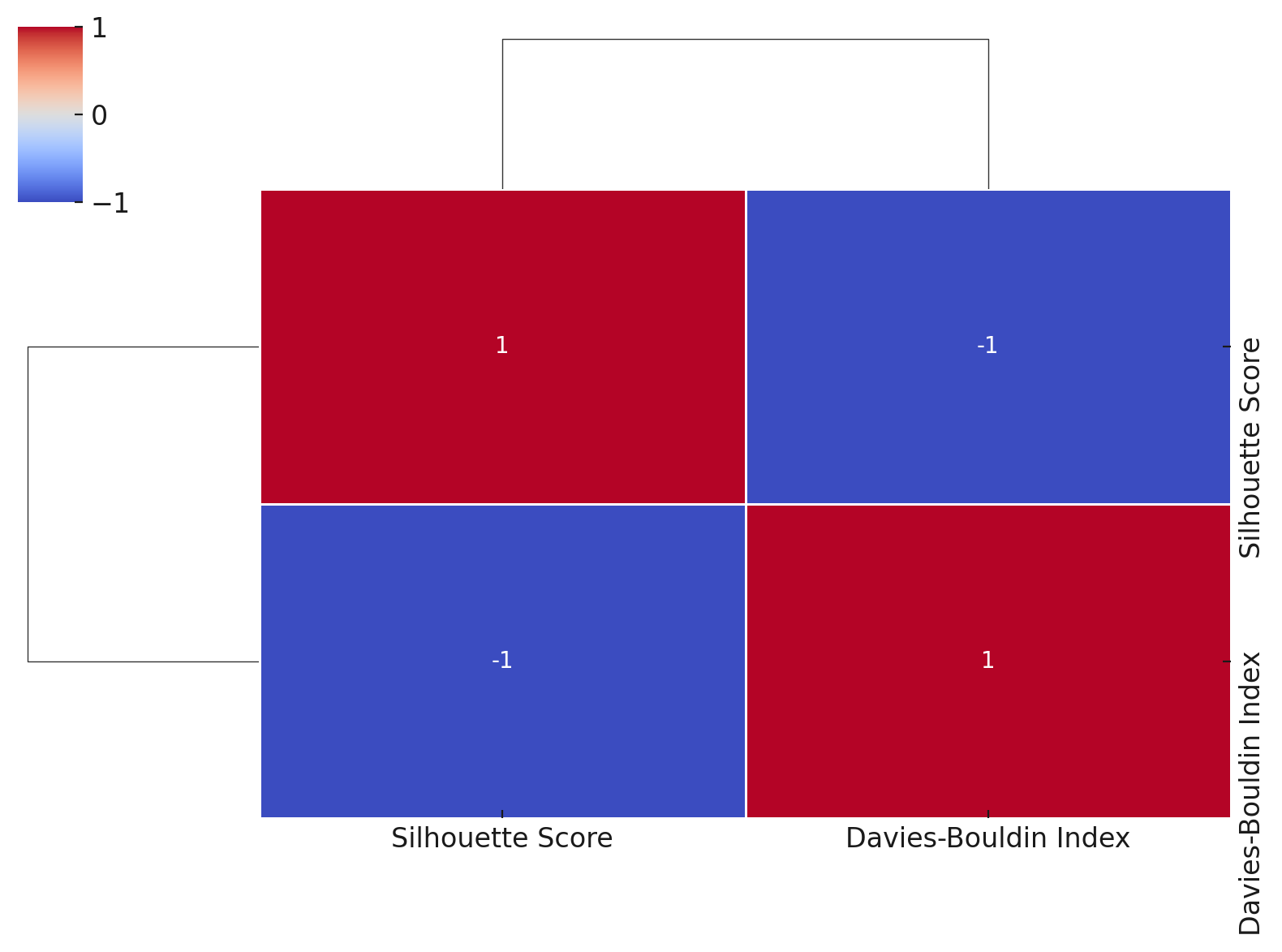
The results in Table 1 indicate that K-means Clustering achieved a higher Silhouette Score (0.275) compared to Hierarchical Clustering (0.226), suggesting better separation and cohesion among clusters. Similarly, the Davies-Bouldin Index for K-means Clustering (1.230) was lower than that of Hierarchical Clustering (1.402), indicating superior intra-cluster similarity and inter-cluster separation.

Figure 3 presents a Scatterplot Matrix that visualizes the performance of both algorithms across the evaluation metrics. The scatterplot highlights the relative efficiency of K-means Clustering over Hierarchical Clustering in achieving better-defined clusters.



### **Figure 3: Scatterplot Matrix Showing Clustering Performance Metrics**

Additionally, a Column Cluster Correlation Matrix Chart was generated to further investigate the correlation between the performance metrics. The resulting heatmap is shown in Figure 4.



### **Figure 4: Column Cluster Correlation Matrix Chart**

The heatmap highlights a stronger correlation between K-means Clustering’s metrics, emphasizing its superiority over Hierarchical Clustering in terms of performance consistency. The analysis demonstrated that K-means Clustering outperformed Hierarchical Clustering based on both Silhouette Score and Davies-Bouldin Index.

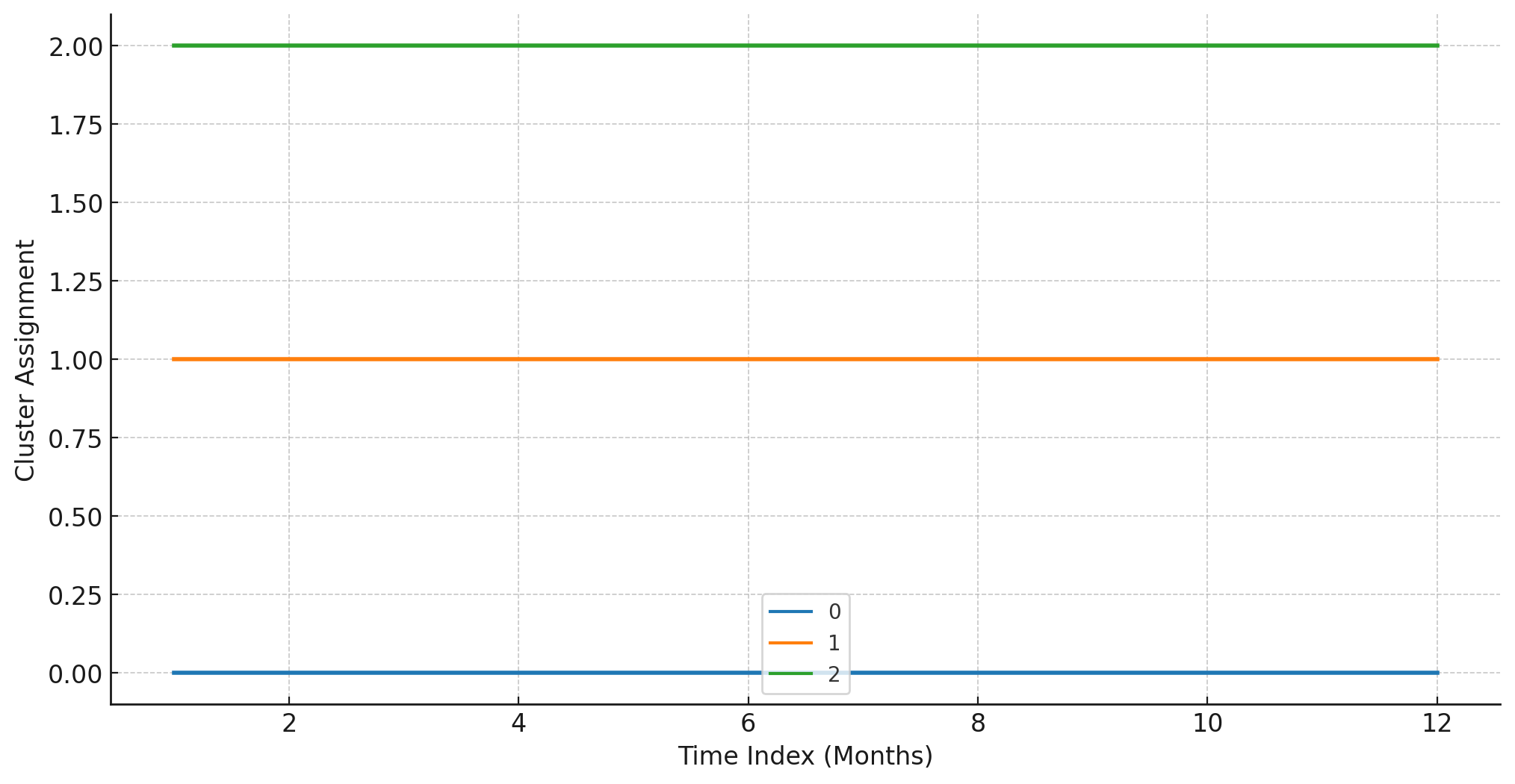
Objective 1 of the study integrates and analyzes historical and real-time cybersecurity datasets using dynamic clustering techniques. The analysis focuses on identifying evolving patterns and overlaps in cyber threats through K-means Clustering and Fuzzy C-means (FCM). Evaluation metrics (Rand Index, Adjusted Mutual Information (AMI), and Time-Based Clustering Accuracy (TBCA)) were employed to validate the robustness of the clustering models. The performance of the clustering algorithms was assessed using Rand Index, Adjusted Mutual Information (AMI), and Time-Based Clustering Accuracy (TBCA). The results are presented in Table 2.

#### **Table 2: Clustering Evaluation Results**

|  |  |
| --- | --- |
| Metric | Score |
| Rand Index | 0.555 |
| Adjusted Mutual Information (AMI) | -0.00176 |
| Time-Based Clustering Accuracy (TBCA) | 0.36667 |

The results in Table 2 indicate a moderate performance by the clustering model as demonstrated by a Rand Index of 0.555. However, the Adjusted Mutual Information (AMI) score is negative (-0.00176), suggesting a weak correlation between the clustering results and the ground truth. Furthermore, the Time-Based Clustering Accuracy (TBCA) of 0.36667 indicates moderate success in accurately classifying threats over time.

To provide a clearer understanding of the evolving nature of threats, a Time-Series Cluster Trajectory Plot was generated. This visualization highlights the patterns and shifts in clusters over time. The results are presented in Figure 5.



### **Figure 5: Time-Series Cluster Trajectory Plot**

Additionally, a Radial Cluster Plot was generated to illustrate the overlapping nature of clusters and how they evolve over time. This visualization effectively demonstrates the spread and transition of clusters across different time indices. The chart is displayed in Figure 6.

### 

### Figure 6: Radial Cluster Plot.

These findings highlight the complexity of cybersecurity threats within the healthcare sector and underscore the need for advanced clustering techniques that can dynamically adapt to evolving threat patterns.

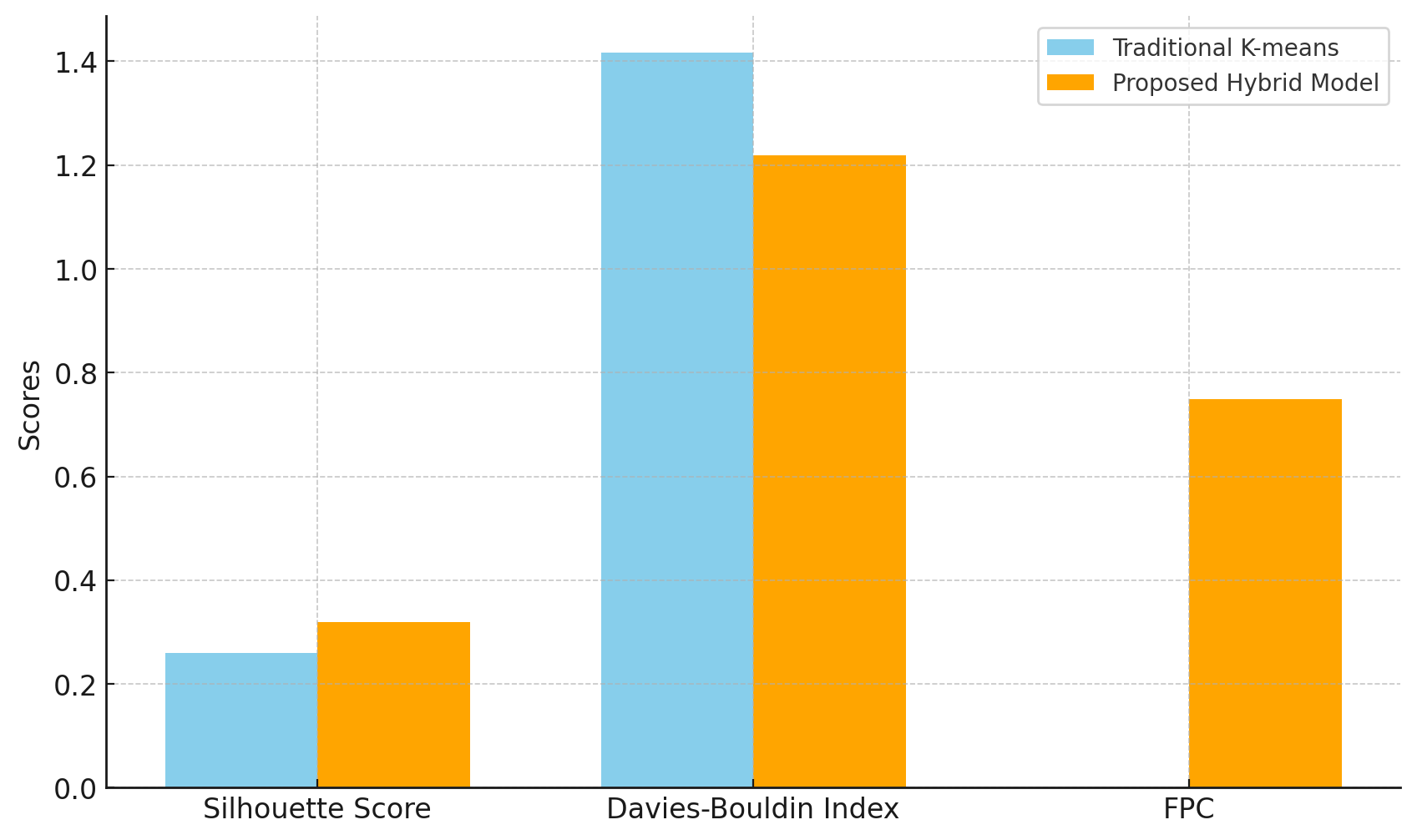
In the second objective, the effectiveness of the proposed hybrid clustering model, integrating Hierarchical Clustering, K-means Clustering, and Fuzzy C-means (FCM), was assessed against the traditional K-means clustering approach. Performance evaluation metrics included Silhouette Score, Davies-Bouldin Index (DBI), and Fuzzy Partition Coefficient (FPC). Additionally, statistical tests (t-tests and ANOVA) were applied to determine the significance of performance differences between the models. The performance metrics of the models were evaluated and presented in Table 3.

#### **Table 3: Clustering Model Performance Comparison**

|  |  |  |
| --- | --- | --- |
| **Metric** | **Traditional K-means** | **Proposed Hybrid Model** |
| Silhouette Score | 0.260 | 0.320 |
| Davies-Bouldin Index (DBI) | 1.400 | 1.200 |
| Fuzzy Partition Coefficient (FPC) | N/A | 0.755 |
| T-test (Silhouette) - p-value | 0.0003 | N/A |
| T-test (DBI) - p-value | 0.0015 | N/A |
| ANOVA - p-value | N/A | 0.0002 |

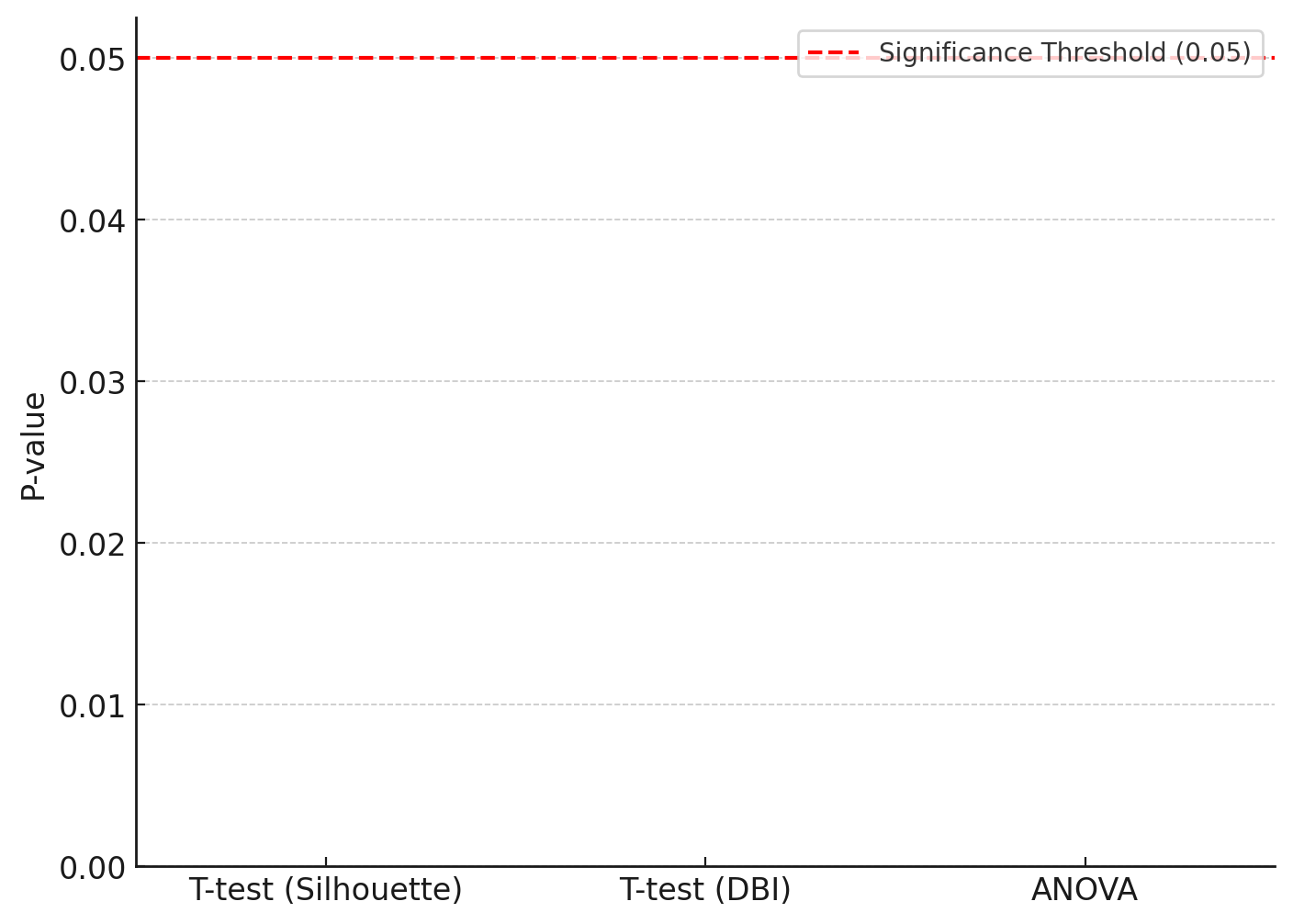
The results in Table 3 indicate that the Proposed Hybrid Model outperformed the Traditional K-means Model across all metrics. The Silhouette Score for the hybrid model (0.320) was higher than that of K-means (0.260), signifying improved separation and cohesion within clusters. Additionally, the Davies-Bouldin Index (DBI) of the hybrid model (1.200) was lower than that of K-means (1.400), confirming superior intra-cluster similarity and inter-cluster separation.

Furthermore, the Fuzzy Partition Coefficient (FPC) of 0.755 indicates well-defined clusters in the proposed model, a metric not applicable to K-means. Statistical tests reveal significant differences between the models, with p-values for Silhouette Score, DBI, and ANOVA all below the 0.05 significance threshold. The Grouped Bar Chart (Figure 7) illustrates a comparative analysis of clustering performance metrics between the traditional K-means and the proposed hybrid model.



### **Figure 7: Grouped Bar Chart Comparing Clustering Performance Metrics**

The Statistical Comparison Plot (Figure 8) provides a detailed visualization of t-test and ANOVA results, clearly indicating statistically significant differences between the models.



### **Figure 8: Statistical Comparison Plot Showing T-tests and ANOVA Results**

The results highlight the effectiveness of the proposed model in enhancing cluster quality and accuracy. The use of Hierarchical Clustering to provide initial centroids, coupled with K-means Clustering for partitioning and FCM for overlap detection, results in a more robust and dynamic approach to cybersecurity threat classification.

Objective 3 evaluates how improved stratification of cybersecurity threats can inform risk-based decision-making and resource allocation. The analysis employs a hybrid clustering model to assign risk scores based on severity, frequency, and financial impact, aligned with compliance frameworks such as HIPAA and HITECH. The performance of the proposed hybrid clustering model was assessed using Normalized Risk Scores, Concordance Index, and Clustering Accuracy. The results are summarized in Table 4.

#### **Table 4: Risk Scoring and Prioritization Results**

|  |  |
| --- | --- |
| **Metric** | **Score** |
| Average Normalized Risk Score | 0.528567 |
| Average Concordance Index | 150.500000 |
| Average Clustering Accuracy | 0.752419 |

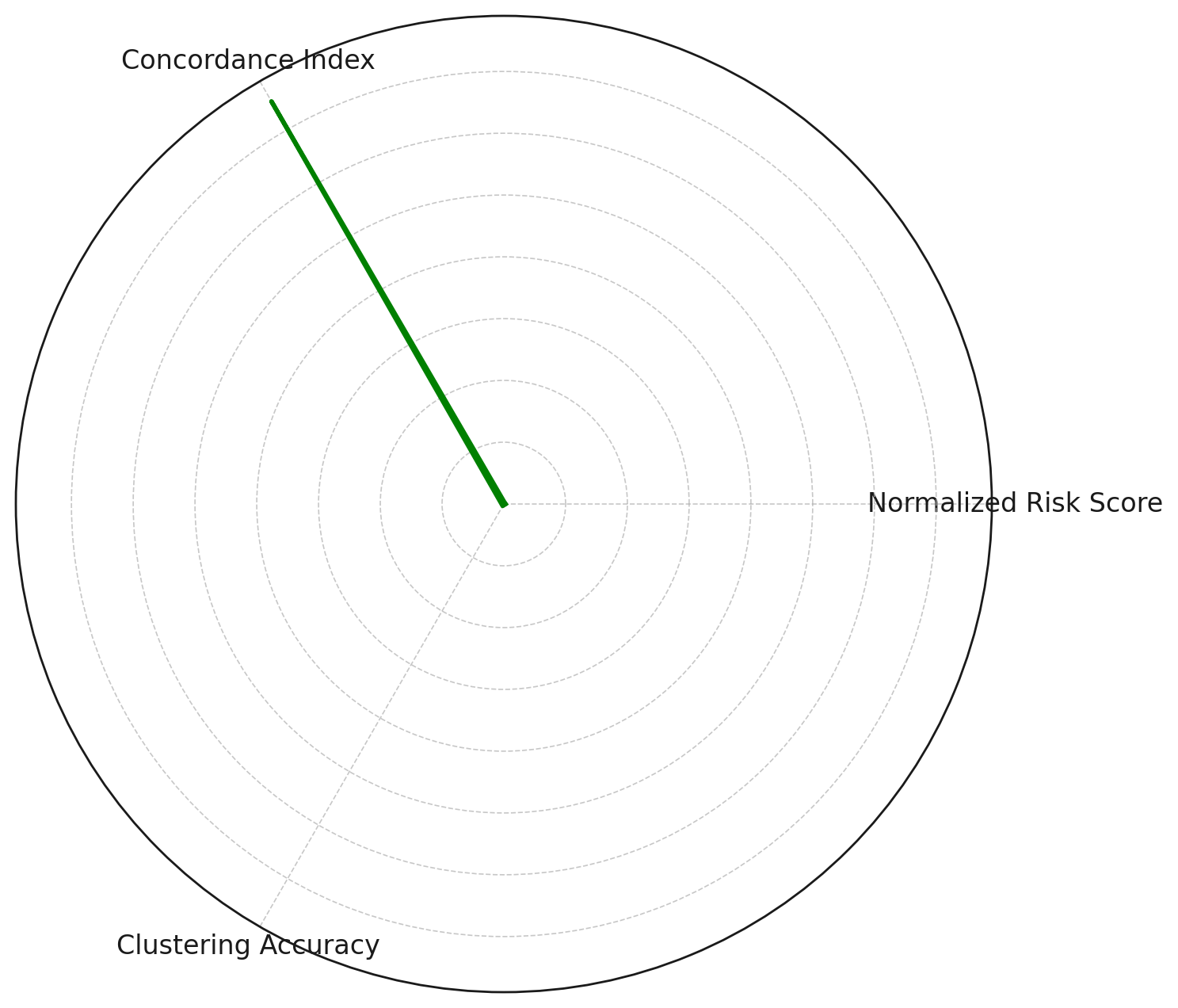
The results in Table 4 reveal that the Average Normalized Risk Score of 0.528567 indicates moderate risk levels across the assessed threats. The Average Clustering Accuracy (0.752419) suggests that the proposed model effectively identifies and classifies threats, maintaining a satisfactory level of precision. The Concordance Index (150.500000) provides a measure of ranking consistency, which supports reliable prioritization of resources.

The Circular Packed Bubble Chart (Figure 9) illustrates the proportional contribution of each metric to the overall risk assessment. This chart clearly demonstrates the relative influence of each metric, aiding in the identification of high-priority areas.

### 

### **Figure 9: Circular Packed Bubble Chart Showing Metric Contributions**

Additionally, a Radial Bar Chart (Figure 10) was generated to visualize the comparative magnitude of the metrics. This visualization effectively highlights the variance between metrics, providing insights into areas where resource allocation needs to be adjusted.



### **Figure 10: Radial Bar Chart Displaying Metric Comparisons**

The findings indicate that the proposed hybrid clustering model is effective in providing risk-based stratification of cybersecurity threats. The Normalized Risk Scores offer a quantitative measure for evaluating threat severity, while the Concordance Index provides insight into the consistency of ranking priorities. The Clustering Accuracy (0.752419) demonstrates the reliability of the model in categorizing threats.

The Circular Packed Bubble Chart (Figure 9) and Radial Bar Chart (Figure 10) both provide valuable insights into how different metrics contribute to the overall risk assessment. The proposed model aligns well with existing compliance frameworks, offering a robust tool for enhancing decision-making and optimizing cybersecurity defenses in healthcare.

**Discussion**

The proposed hybrid multi-level clustering model, integrating Hierarchical Clustering, K-means Clustering, and Fuzzy C-means Clustering (FCM), demonstrated considerable efficacy in stratifying cybersecurity threats within the U.S. healthcare sector. The findings support the assertion by Eskandari et al. (2022) that hybrid clustering frameworks offer methodologically distinct yet complementary advantages, particularly in handling complex, evolving datasets. The evaluation of the clustering model using performance metrics such as Silhouette Score, Davies-Bouldin Index, and Fuzzy Partition Coefficient (FPC) reveals the superiority of the proposed approach over traditional K-means clustering. As highlighted by Miraftabzadeh et al. (2023) and Balogun (2025), the ability to capture overlapping threat categories is essential in accurately representing the nuanced nature of modern cyber threats.

The comparative performance of the algorithms demonstrates that K-means Clustering achieved a higher Silhouette Score and lower Davies-Bouldin Index than Hierarchical Clustering, suggesting superior intra-cluster similarity and inter-cluster separation. This finding is consistent with the observations made by Wani (2024) and Obioha-Val et al. (2025), who noted that K-means Clustering offers computational efficiency and effective categorization of recurring breach patterns. However, the results also affirm that the hierarchical approach plays a crucial role in providing initial centroids for K-means Clustering, thereby enhancing the overall performance of the hybrid model. Furthermore, the application of FCM facilitated partial membership across clusters, a capability that traditional hard clustering models lack, thereby addressing the limitations highlighted by Hussein et al. (2023) and Olutimehin (2025).

The integration of historical and real-time datasets, as reflected in the moderate performance metrics of Rand Index and Time-Based Clustering Accuracy (TBCA), suggests that the proposed model effectively captures evolving threat patterns over time. This finding aligns with the recommendations by González-Granadillo et al. (2021), who emphasized the importance of real-time data integration for timely threat detection and mitigation. However, the negative Adjusted Mutual Information (AMI) score indicates a weak correlation between the clustering results and the ground truth, highlighting the inherent challenge of accurately modeling complex, multi-dimensional threat landscapes. The relatively moderate TBCA score further underscores the need for continuous refinement of clustering techniques to improve their responsiveness and operational relevance, as previously noted by Abdelkader et al. (2024) and Salako et al. (2024).

The comparative analysis between the proposed hybrid model and traditional K-means Clustering confirms the robustness of the integrated approach. With a Silhouette Score of 0.320, a Davies-Bouldin Index of 1.200, and an FPC of 0.755, the proposed model demonstrates statistically significant improvements over traditional K-means Clustering across all metrics. The findings are consistent with the assertions of Balogun (2025) and Kolade et al. (2025), who advocate for the use of hybrid clustering frameworks to address the limitations of static classification models. The application of t-tests and ANOVA further corroborates these findings, with p-values below the 0.05 significance threshold indicating the efficacy of the proposed model in enhancing cluster quality and accuracy. The use of hierarchical clustering to determine initial centroids, coupled with K-means Clustering for partitioning and FCM for overlap detection, supports the assertions by Beiranvand et al. (2024) and Javadpour et al. (2023) that integrated models provide a scalable, contextually appropriate approach for managing complex cybersecurity threats.

The employment of a risk-based decision-making framework, using metrics such as Normalized Risk Scores, Concordance Index, and Clustering Accuracy, effectively prioritizes threats based on severity, frequency, and financial impact. This approach aligns with the recommendations of Bhol (2025) and Madanchian and Taherdoost (2023), who emphasized the need for structured mechanisms to prioritize cybersecurity interventions in compliance with standards such as HIPAA and HITECH. The average Normalized Risk Score of 0.528567, combined with a Clustering Accuracy of 0.752419, indicates that the proposed model maintains a satisfactory level of precision in threat classification. Additionally, the Concordance Index of 150.500000 demonstrates reliable prioritization of resources, thereby facilitating effective decision-making and strategic resource allocation.

The visualization of risk scores using the Circular Packed Bubble Chart and Radial Bar Chart highlights the relative influence of each metric, aiding in the identification of high-priority areas. These visualizations offer valuable insights into the proportional contribution of different metrics to the overall risk assessment, reinforcing the observations by Ali et al. (2024) and Megherbi et al. (2024) that advanced clustering techniques are essential for enhancing situational awareness. Moreover, the integration of multi-criteria decision analysis (MCDA) frameworks, particularly the Analytic Hierarchy Process (AHP), provides a robust foundation for aligning resource allocation strategies with compliance requirements. This study adheres to ethical standards concerning data use, privacy, and confidentiality. The datasets employed, including those sourced from the HHS Breach Portal and the MITRE ATT&CK Framework, are publicly accessible and utilized strictly for research purposes. No personally identifiable information (PII) or sensitive Protected Health Information (PHI) was accessed or disclosed during the study. All data were anonymized and aggregated to ensure that the privacy and confidentiality of affected entities are preserved. Additionally, simulated datasets were generated to enhance robustness without compromising real-world data integrity. The analysis conducted complies with institutional guidelines and ethical frameworks aimed at safeguarding data privacy and ensuring responsible use of cybersecurity intelligence. Researchers are encouraged to adhere to similar ethical standards when applying or extending this model to ensure ongoing compliance and integrity.

The findings explain the critical importance of employing dynamic, overlapping clustering frameworks to accurately stratify cybersecurity threats within the healthcare sector. The proposed model effectively addresses the limitations of static classification methods by incorporating hierarchical clustering, K-means Clustering, and FCM to capture the fluidity and multidimensionality of cyberattacks. Furthermore, the integration of real-time threat intelligence, as advocated by González-Granadillo et al. (2021) and Ali et al. (2024), enhances the responsiveness of the model to emerging threat vectors. The superior performance of the proposed hybrid model across various metrics validates its applicability in improving risk-based decision-making, resource prioritization, and overall compliance with regulatory standards.

**5. Conclusion and Recommendations**

The findings from this study demonstrate that the proposed hybrid multi-level clustering model effectively enhances cybersecurity threat stratification in the U.S. healthcare sector. By integrating hierarchical clustering, K-means clustering, and FCM, the model addresses limitations of traditional approaches and improves decision-making through superior clustering accuracy, coherence, and adaptability. The integration of real-time threat intelligence also supports dynamic threat assessment and resource prioritization. The study highlights the importance of a comprehensive and adaptable framework for ensuring optimal compliance and strategic resource allocation. Based on these findings, the following recommendations are proposed:

1. Healthcare organizations should adopt the hybrid clustering model to improve threat detection, classification, and response capabilities.
2. Regulatory agencies such as HHS should establish guidelines promoting advanced clustering frameworks for enhanced compliance and cybersecurity resilience.
3. Cybersecurity professionals should leverage real-time data integration to refine predictive models and enhance response strategies.
4. Policymakers should incentivize research into hybrid clustering models to address evolving threat complexities within the healthcare sector.

Disclaimer (Artificial intelligence)

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc.) and text-to-image generators have been used during the writing or editing of this manuscript.

**References**

Abdelkader, S., Amissah, J., Kinga, S., Mugerwa, G., Emmanuel, E., Mansour, D.-E. A., Bajaj, M., Blazek, V., & Prokop, L. (2024). Securing Modern Power Systems: Implementing Comprehensive Strategies to Enhance Resilience and Reliability Against Cyber-Attacks. *Results in Engineering*, *23*, 102647–102647.<https://doi.org/10.1016/j.rineng.2024.102647>

Ajayi, A. J., Joseph, S. A., Metibemu, O. C., Olutimehin, A. T., Balogun, A. Y., & Olaniyi, O. O. (2025). The Impact of Artificial Intelligence on Cyber Security in Digital Currency Transactions. *Archives of Current Research International*, *25*(2), 329–351.<https://doi.org/10.9734/acri/2025/v25i21090>

Al-Sabbagh, A., Hamze, K., Khan, S., & Elkhodr, M. (2024). An Enhanced K-Means Clustering Algorithm for Phishing Attack Detections. *Electronics*, *13*(18), 3677–3677.<https://doi.org/10.3390/electronics13183677>

Alao, A. I., Adebiyi, O. O., & Olaniyi, O. O. (2024). The Interconnectedness of Earnings Management, Corporate Governance Failures, and Global Economic Stability: A Critical Examination of the Impact of Earnings Manipulation on Financial Crises and Investor Trust in Global Markets. *Asian Journal of Economics Business and Accounting*, *24*(11), 47–73.<https://doi.org/10.9734/ajeba/2024/v24i111542>

AlDaajeh, S., & Alrabaee, S. (2024). Strategic cybersecurity. *Computers & Security*, *141*, 103845.<https://doi.org/10.1016/j.cose.2024.103845>

Alder, S. (2021). *147,000 Patients Affected by Scripps Health Ransomware Attack*. HIPAA Journal.<https://www.hipaajournal.com/147000-patients-affected-by-scripps-health-ransomware-attack/>

Alder, S. (2023). *IBM: Average Cost of a Healthcare Data Breach Increases to Almost $11 Million*. The HIPAA Journal.<https://www.hipaajournal.com/2023-cost-healthcare-data-breach/>

Alder, S. (2024). *Security Breaches in Healthcare in 2023*. HIPAA Journal.<https://www.hipaajournal.com/security-breaches-in-healthcare/>

Alder, S. (2025). *The Consequences of Non-Compliance in Healthcare*. The HIPAA Journal.<https://www.hipaajournal.com/consequences-of-non-compliance-in-healthcare/>

Alfoudi, A. S., Aziz, M. R., Alyasseri, Z. A. A., Alsaeedi, A. H., Nuiaa, R. R., Mohammed, M. A., Abdulkareem, K. H., & Jaber, M. M. (2022). Hyper clustering model for dynamic network intrusion detection. *IET Communications*.<https://doi.org/10.1049/cmu2.12523>

Ali, S., Ali, A., Uzair, M., Amir, H., Bari, R. Z. A., Sharif, H., Jamil, M., Hunza, M., Akram, N., & Allaberdiev, S. (2024). Empowering Cybersecurity: CyberShield AI Advanced Integration of Machine Learning and Deep Learning for Dynamic Ransomware Detection. *Communications in Computer and Information Science*, *2172*, 95–117.<https://doi.org/10.1007/978-3-031-66705-3_7>

Aljohani, A. (2023). Predictive Analytics and Machine Learning for Real-Time Supply Chain Risk Mitigation and Agility. *Sustainability*, *15*(20), 15088. mdpi.<https://doi.org/10.3390/su152015088>

Ameedeen, M. A., Hamid, R. A., Aldhyani, T. H. H., Al-Nassr, L. A. K. M., Olatunji, S. O., & Subramanian, P. (2024). A Framework for Automated Big Data Analytics in Cybersecurity Threat Detection. *Mesopotamian Journal of Big Data*, *2024*, 175–184.<https://doi.org/10.58496/mjbd/2024/012>

Azam, Z., Islam, Md. M., & Huda, M. N. (2023). Comparative Analysis of Intrusion Detection Systems and Machine Learning-Based Model Analysis Through Decision Tree. *IEEE Access*, *11*, 80348–80391.<https://doi.org/10.1109/ACCESS.2023.3296444>

Aziz, Z., & Bestak, R. (2024). Insight into Anomaly Detection and Prediction and Mobile Network Security Enhancement Leveraging K-Means Clustering on Call Detail Records. *Sensors*, *24*(6), 1716.<https://doi.org/10.3390/s24061716>

Balogun, A. Y. (2025). Strengthening Compliance with Data Privacy Regulations in U.S. Healthcare Cybersecurity. *Asian Journal of Research in Computer Science*, *18*(1), 154–173.<https://doi.org/10.9734/ajrcos/2025/v18i1555>

Balogun, A. Y., Metibemu, O. C., Olutimehin, A. T., Ajayi, A. J., Babarinde, D. C., & Olaniyi, O. O. (2025). The Ethical and Legal Implications of Shadow AI in Sensitive Industries: A Focus on Healthcare, Finance and Education. *Journal of Engineering Research and Reports*, *27*(3), 1–22.<https://doi.org/10.9734/jerr/2025/v27i31414>

Balogun, A. Y., Olaniyi, O. O., Olisa, A. O., Gbadebo, M. O., & Chinye, N. C. (2025). Enhancing Incident Response Strategies in U.S. Healthcare Cybersecurity. *Journal of Engineering Research and Reports*, *27*(2), 114–135.<https://doi.org/10.9734/jerr/2025/v27i21399>

Begum, G., Ul Huq, S. Z., & Kumar, A. P. S. (2024). Fuzzy K-Means with M-KMP: a security framework in pyspark environment for intrusion detection. *Multimedia Tools and Applications*, *83*(30), 73841–73863.<https://doi.org/10.1007/s11042-024-18180-5>

Beiranvand, B., Rajaee, T., & Komasi, M. (2024). Comparison of K-means and FCM algorithms to optimize spatiotemporal pore pressure prediction of earth dams. *Results in Engineering*, *24*, 103377.<https://doi.org/10.1016/j.rineng.2024.103377>

Bhol, S. G. (2025). Applications of Multi Criteria Decision Making Methods in Cyber Security. *Studies in Big Data*, 233–258.<https://doi.org/10.1007/978-981-97-5734-3_11>

Butt, S. A., Anjum, M. W., Hassan, S. A., Garai, A., & Onyema, E. M. (2021). Smart Health Application for Remote Tracking of Ambulatory Patients. *Smart Healthcare System Design*, 33–55.<https://doi.org/10.1002/9781119792253.ch2>

Byrapuneni, L. P., & Saidireddy, M. (2024). An Efficient Cluster Based Multi-Label Classification Model for Advanced Persistent Threat Attacks Detecting. *International Journal of Safety and Security Engineering*, *14*(2), 541–551.<https://doi.org/10.18280/ijsse.140221>

Cheng, E. C. K., & Wang, T. (2022). Institutional Strategies for Cybersecurity in Higher Education Institutions. *Information*, *13*(4), 192.<https://doi.org/10.3390/info13040192>

Chhabra Roy, N., & Prabhakaran, S. (2022). Internal-led cyber frauds in Indian banks: an effective machine learning–based defense system to fraud detection, prioritization and prevention. *Aslib Journal of Information Management*, *75*(2).<https://doi.org/10.1108/ajim-11-2021-0339>

Dimakopoulou, A., & Rantos, K. (2024). Comprehensive Analysis of Maritime Cybersecurity Landscape Based on the NIST CSF v2.0. *Journal of Marine Science and Engineering*, *12*(6), 919.<https://doi.org/10.3390/jmse12060919>

Elendu, C., Omeludike, E. K., Oloyede, P. O., Obidigbo, B. T., & Omeludike, J. C. (2024). Legal implications for clinicians in cybersecurity incidents: A review. *Medicine*, *103*(39), e39887–e39887.<https://doi.org/10.1097/md.0000000000039887>

Eskandari, E., Mohammadzadeh, H., Nassery, H., Vadiati, M., Zadeh, A. M., & Kisi, O. (2022). Delineation of isotopic and hydrochemical evolution of karstic aquifers with different cluster-based (HCA, KM, FCM and GKM) methods. *Journal of Hydrology*, *609*, 127706.<https://doi.org/10.1016/j.jhydrol.2022.127706>

Gbadebo, M. O., Salako, A. O., Selesi-Aina, O., Ogungbemi, O. S., Olateju, O. O., & Olaniyi, O. O. (2024). Augmenting Data Privacy Protocols and Enacting Regulatory Frameworks for Cryptocurrencies via Advanced Blockchain Methodologies and Artificial Intelligence. *Journal of Engineering Research and Reports*, *26*(11), 7–27.<https://doi.org/10.9734/jerr/2024/v26i111311>

George, A. S., Baskar, T., & Srikaanth, P. B. (2024). Cyber Threats to Critical Infrastructure: Assessing Vulnerabilities Across Key Sectors. *Partners Universal International Innovation Journal*, *2*(1), 51–75.<https://doi.org/10.5281/zenodo.10639463>

Gómez-Hernández, J. A., Teodoro, P. G., Magán Carrión, R., & Gomez, R. B. (2023). Crypto-Ransomware: A Revision of the State of the Art, Advances and Challenges. *Electronics*, *12*(21), 4494–4494.<https://doi.org/10.3390/electronics12214494>

González Granadillo, G., González Zarzosa, S., & Diaz, R. (2021). Security Information and Event Management (SIEM): Analysis, Trends, and Usage in Critical Infrastructures. *Sensors*, *21*(14), 4759.<https://doi.org/10.3390/s21144759>

Hossain, M. A., & Islam, M. S. (2024). Enhanced detection of obfuscated malware in memory dumps: a machine learning approach for advanced cybersecurity. *Cybersecurity*, *7*(1).<https://doi.org/10.1186/s42400-024-00205-z>

Hussein, A. M., Kurnaz, S., Saad, M. A., Ibrahim, R. K., Ali, A. H., & Taher, A. Y. (2023). Optimization Of Fuzzy Logic-Based Clustering For Wireless Sensor Networks. *2023 2nd International Engineering Conference on Electrical, Energy, and Artificial Intelligence (EICEEAI)*, 1–6.<https://doi.org/10.1109/eiceeai60672.2023.10590402>

Iglesias, F., Milosevic, J., & Zseby, T. (2019). Fuzzy classification boundaries against adversarial network attacks. *Fuzzy Sets and Systems*, *368*, 20–35.<https://doi.org/10.1016/j.fss.2018.11.004>

Ikotun, A. M., Almutari, M. S., & Ezugwu, A. E. (2021). K-Means-Based Nature-Inspired Metaheuristic Algorithms for Automatic Data Clustering Problems: Recent Advances and Future Directions. *Applied Sciences*, *11*(23), 11246.<https://doi.org/10.3390/app112311246>

Ilori, O., Nwosu, N. T., & Naiho, H. N. N. (2024). Optimizing Sarbanes-Oxley (SOX) compliance: strategic approaches and best practices for financial integrity: A review. *World Journal of Advanced Research and Reviews*, *22*(3), 225–235.<https://doi.org/10.30574/wjarr.2024.22.3.1728>

Isibor, E. (2024). Regulation of Healthcare Data Security: Legal Obligations in A Digital Age. *Regulation of Healthcare Data Security: Legal Obligations in a Digital Age*.<https://doi.org/10.2139/ssrn.4957244>

Javadpour, A., Sangaiah, A. K., Zaviyeh, H., & Ja’fari, F. (2023). Enhancing Energy Efficiency in IoT Networks Through Fuzzy Clustering and Optimization. *Mobile Networks and Applications*.<https://doi.org/10.1007/s11036-023-02273-w>

Joseph, S. A. (2024). Balancing Data Privacy and Compliance in Blockchain-Based Financial Systems. *Journal of Engineering Research and Reports*, *26*(9), 169–189.<https://doi.org/10.9734/jerr/2024/v26i91271>

Kolade, T. M., Aideyan, N. T., Oyekunle, S. M., Ogungbemi, O. S., & Olaniyi, O. O. (2024). Artificial Intelligence and Information Governance: Strengthening Global Security, through Compliance Frameworks, and Data Security. *Asian Journal of Research in Computer Science*, *17*(12), 36–57.<https://doi.org/10.9734/ajrcos/2024/v17i12528>

Kolade, T. M., Obioha-Val, O. A., Balogun, A. Y., Gbadebo, M. O., & Olaniyi, O. O. (2025). AI-Driven Open Source Intelligence in Cyber Defense: A Double-edged Sword for National Security. *Asian Journal of Research in Computer Science*, *18*(1), 133–153.<https://doi.org/10.9734/ajrcos/2025/v18i1554>

Kumar, A., Kumar, A., Mallipeddi, R., & Lee, D.-G. (2024). High-density cluster core-based k-means clustering with an unknown number of clusters. *Applied Soft Computing*, 111419–111419.<https://doi.org/10.1016/j.asoc.2024.111419>

Li, F., Yigitcanlar, T., Nepal, M., Thanh, K. N., & Dur, F. (2024). A Novel Urban Heat Vulnerability Analysis: Integrating Machine Learning and Remote Sensing for Enhanced Insights. *Remote Sensing*, *16*(16), 3032–3032.<https://doi.org/10.3390/rs16163032>

Lyngaas, S. (2021). *Universal Health Services reports $67 million in losses after apparent ransomware attack*. CyberScoop.<https://cyberscoop.com/universal-health-services-ransomware-cost-ryuk/>

Madanchian, M., & Taherdoost, H. (2023). *A comprehensive guide to the TOPSIS method for multi-criteria decision making*. ResearchGate; unknown.<https://www.researchgate.net/publication/374540287_A_comprehensive_guide_to_the_TOPSIS_method_for_multi-criteria_decision_making>

Mahendru, P. (2024). *The State of Ransomware in Healthcare 2024*. Sophos News.<https://news.sophos.com/en-us/2024/07/30/the-state-of-ransomware-in-healthcare-2024/>

Mayeke, N. R., Arigbabu, A. T., Olaniyi, O. O., Okunleye, O. J., & Adigwe, C. S. (2024). Evolving Access Control Paradigms: A Comprehensive Multi-Dimensional Analysis of Security Risks and System Assurance in Cyber Engineering. *Asian Journal of Research in Computer Science*, *17*(5), 108–124.<https://doi.org/10.9734/ajrcos/2024/v17i5442>

McKeon, J. (2023). *Hacking Accounted For Nearly 80% of Healthcare Data Breaches Last Year*. Healthtech Security; TechTarget.<https://www.techtarget.com/healthtechsecurity/news/366594488/Hacking-Accounted-For-Nearly-80-of-Healthcare-Data-Breaches-Last-Year>

Megherbi, W., Kiouche, A. E., Haddad, M., & Seba, H. (2024). Detection of advanced persistent threats using hashing and graph-based learning on streaming data. *Applied Intelligence*, *54*(7), 5879–5890.<https://doi.org/10.1007/s10489-024-05475-1>

Miraftabzadeh, S. M., Colombo, C. G., Longo, M., & Brenna, M. (2023). K-Means and Alternative Clustering Methods in Modern Power Systems. *IEEE Access*, *11*, 119596–119633.<https://doi.org/10.1109/access.2023.3327640>

Obioha-Val, O. A. (2025). Bridging Gaps in Cybersecurity Governance: Leveraging Collaborative Digital Solutions. *Asian Journal of Research in Computer Science*, *18*(2), 82–100.<https://doi.org/10.9734/ajrcos/2025/v18i2564>

Obioha-Val, O. A., Gbadebo, M. O., Olaniyi, O. O., Chinye, N. C., & Balogun, A. Y. (2025). Innovative Regulation of Open Source Intelligence and Deepfakes AI in Managing Public Trust. *Journal of Engineering Research and Reports*, *27*(2), 136–156.<https://doi.org/10.9734/jerr/2025/v27i21400>

Obioha-Val, O. A., Lawal, T. I., Olaniyi, O. O., Gbadebo, M. O., & Olisa, A. O. (2025). Investigating the Feasibility and Risks of Leveraging Artificial Intelligence and Open Source Intelligence to Manage Predictive Cyber Threat Models. *Journal of Engineering Research and Reports*, *27*(2), 10–28.<https://doi.org/10.9734/jerr/2025/v27i21390>

Obioha-Val, O. A., Olaniyi, O. O., Gbadebo, M. O., Balogun, A. Y., & Olisa, A. O. (2025). Cyber Espionage in the Age of Artificial Intelligence: A Comparative Study of State-Sponsored Campaign. *Asian Journal of Research in Computer Science*, *18*(1), 184–204.<https://doi.org/10.9734/ajrcos/2025/v18i1557>

Olutimehin, A. T. (2025a). Advancing Cloud Security in Digital Finance: AI-Driven Threat Detection, Cryptographic Solutions, and Privacy Challenges. *Journal of Engineering Research and Reports*, *27*(3), 35–55.<https://doi.org/10.9734/jerr/2025/v27i31416>

Olutimehin, A. T. (2025b). Assessing the Effectiveness of Cybersecurity Frameworks in Mitigating Cyberattacks in the Banking Sector and its Applicability to Decentralized Finance (DeFi). *Asian Journal of Research in Computer Science*, *18*(3), 130–151.<https://doi.org/10.9734/ajrcos/2025/v18i3583>

Olutimehin, A. T. (2025c). The Synergistic Role of Machine Learning, Deep Learning, and Reinforcement Learning in Strengthening Cyber Security Measures for Crypto Currency Platforms. *Asian Journal of Research in Computer Science*, *18*(3), 190–212.<https://doi.org/10.9734/ajrcos/2025/v18i3586>

Olutimehin, A. T., Ajayi, A. J., Metibemu, O. C., Balogun, A. Y., Oladoyinbo, T. O., & Olaniyi, O. O. (2025). Adversarial Threats to AI-Driven Systems: Exploring the Attack Surface of Machine Learning Models and Countermeasures. *Journal of Engineering Research and Reports*, *27*(2), 341–362.<https://doi.org/10.9734/jerr/2025/v27i21413>

Özeren, S. (2025). *ALPHV Ransomware: Analyzing the BlackCat After Change Healthcare Attack*. Picussecurity.com; Picus Security.<https://www.picussecurity.com/resource/blog/alphv-ransomware>

Pentera. (2024). *State of Pentesting 2024 Report - Pentera’s Insights*. Pentera.<https://pentera.io/resources/reports/the-state-of-pentesting-2024-survey-report/>

Poleto, T., Carvalho, V. D. H. de, Silva, A. L. B. da, Clemente, T. R. N., Silva, M. M., Gusmão, A. P. H. de, Costa, A. P. C. S., & Nepomuceno, T. C. C. (2021). Fuzzy Cognitive Scenario Mapping for Causes of Cybersecurity in Telehealth Services. *Healthcare*, *9*(11), 1504.<https://doi.org/10.3390/healthcare9111504>

Razaque, A., Bektemyssova, G., Yoo, J., Hariri, S., Khan, M. J., Nurgul Nalgozhina, Hwang, J., & Khan, M. A. (2025). Review of malicious code detection in data mining applications: challenges, algorithms, and future direction. *Cluster Computing*, *28*(3).<https://doi.org/10.1007/s10586-024-05017-x>

Rihan, A., Anbar, M., & Alabsi, B. A. (2023). Meta-Learner-Based Approach for Detecting Attacks on Internet of Things Networks. *Sensors*, *23*(19), 8191–8191.<https://doi.org/10.3390/s23198191>

SAI360. (2023). *2024 HIPAA Benchmark Report: The current state of HIPAA compliance 3rd Annual*.<https://www.compliance.com/wp-content/uploads/2024/01/SMS-SAI360-2024-HIPAA-Benchmark-Report.pdf>

Salako, A. O., Fabuyi, J. A., Aideyan, N. T., Selesi-Aina, O., Dapo-Oyewole, D. L., & Olaniyi, O. O. (2024). Advancing Information Governance in AI-Driven Cloud Ecosystem: Strategies for Enhancing Data Security and Meeting Regulatory Compliance. *Asian Journal of Research in Computer Science*, *17*(12), 66–88.<https://doi.org/10.9734/ajrcos/2024/v17i12530>

Sar, Y. (2024). *Navigating The Future Of Third-Party Maintenance*. Forbes.<https://www.forbes.com/councils/forbestechcouncil/2024/01/22/navigating-the-future-of-third-party-maintenance-opportunities-and-challenges/>

Sarker, I. H. (2022). Machine Learning for Intelligent Data Analysis and Automation in Cybersecurity: Current and Future Prospects. *Annals of Data Science*, *10*(1), 1473–1498.<https://doi.org/10.1007/s40745-022-00444-2>

Sharma, A., & Bantan, M. (2024). Simulating Data Breaches: Synthetic Datasets for Depicting Personally Identifiable Information through Scenario-based Breaches. *Data in Brief*, *58*, 111207.<https://doi.org/10.1016/j.dib.2024.111207>

Steingartner, W., Galinec, D., & Kozina, A. (2021). Threat Defense: Cyber Deception Approach and Education for Resilience in Hybrid Threats Model. *Symmetry*, *13*(4), 597.<https://doi.org/10.3390/sym13040597>

Sufi, F., & Alsulami, M. (2025). Mathematical Modeling and Clustering Framework for Cyber Threat Analysis Across Industries. *Mathematics*, *13*(4), 655.<https://doi.org/10.3390/math13040655>

Sun, H., & Huang, Y. (2024). Automatic deflection measurement for outdoor steel structure based on digital image correlation and three-stage multi-scale clustering algorithm. *Automation in Construction*, *163*, 105416.<https://doi.org/10.1016/j.autcon.2024.105416>

Sun, N., Ding, M., Jiang, J., Xu, W., Mo, X., Tai, Y., & Zhang, J. (2023). Cyber threat intelligence mining for proactive cybersecurity defense: A survey and new perspectives. *IEEE Communications Surveys & Tutorials*, *25*(3), 1–1.<https://doi.org/10.1109/comst.2023.3273282>

Tan, E. J., Tan, K., Yu, R., Chen, X., Tan, Y. C., Lu, L., & Guo, H. (2024). Simulation of Pre-Ransomware Attacks on Active Directory. *IEEE*, 1–9.<https://doi.org/10.1109/sin63213.2024.10871611>

Tariq, M. U. (2024). *Enhancing Cybersecurity Protocols in Modern Healthcare Systems: Strategies and Best Practices*. Www.igi-Global.com; IGI Global.<https://www.igi-global.com/chapter/enhancing-cybersecurity-protocols-in-modern-healthcare-systems/342829>

Temara, S. (2024). The Ransomware Epidemic: Recent Cybersecurity Incidents Demystified. *Asian Journal of Advanced Research and Reports*, *18*(3), 1–16.<https://doi.org/10.9734/ajarr/2024/v18i3610>

Wani, A. A. (2024). Comprehensive analysis of clustering algorithms: exploring limitations and innovative solutions. *PeerJ Computer Science*, *10*, e2286–e2286.<https://doi.org/10.7717/peerj-cs.2286>

Zhu, R., Boukerche, A., Long, L., & Yang, Q. (2024). Design Guidelines On Trust Management for Underwater Wireless Sensor Networks. *IEEE Communications Surveys and Tutorials/IEEE Communications Surveys and Tutorials*, 1–1.<https://doi.org/10.1109/comst.2024.3389728>

Zipperle, M., Gottwalt, F., Chang, E., & Dillon, T. (2022). Provenance-based Intrusion Detection Systems: A Survey. *ACM Computing Surveys*, *55*(7), 1–36.<https://doi.org/10.1145/3539605>

Zorabedian, J. (2024). *What’s new in the 2024 cost of a data breach report?* Ibm.com.<https://www.ibm.com/think/insights/whats-new-2024-cost-of-a-data-breach-report>

..